

CC purposes to treat diseases or disorders including cancer, immune-related
CC disease, cardiovascular disease, brain or neuronal associated disease,
CC metabolic and inflammatory disorders. Accordingly, they exhibit
CC cytostatic, neuroprotective, immunomodulator and antiinflammatory
CC activities. This polypeptide sequence is a human kinase protein sequence
CC of the invention.

XX
SQ Sequence 507 AA;

Query Match 100.0%; Score 2670; DB 8; Length 507;

Best Local Similarity 100.0%; Pred. No. 1,2e-250; Indels 0; Gaps 0;

Matches 507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPEPAPRRAREREREDESEDESDESDILEESPCCGMQKRREQVNGNMPGIQSTFLAMD 60
DB 1 MAPEPAPRRAREREREDESEDESDESDILEESPCCGMQKRREQVNGNMPGIQSTFLAMD 60
QY 61 TEEGVEVWVWNEHLHFGDRKAPAAHEEKIQTFVEQVLVDHPNIVKLHKWLDTSEACAVY 120
DB 61 TEEGVEVWVWNEHLHFGDRKAPAAHEEKIQTFVEQVLVDHPNIVKLHKWLDTSEACAVY 120
QY 121 FITEVSSGSLKQPLKTKKQNHKAMNARAKMCTQIISALSFTHACSPPIIHGULTSDT 180
DB 121 FITEVSSGSLKQPLKTKKQNHKAMNARAKMCTQIISALSFTHACSPPIIHGULTSDT 180
QY 181 IFIQHNGLIKIGSVWHRIFSNALRPPTALPDDLSPIRAEREELRNLFPPPEYGEVADG 240
DB 181 IFIQHNGLIKIGSVWHRIFSNALRPPTALPDDLSPIRAEREELRNLFPPPEYGEVADG 240
QY 241 TAVVDIFSFGMCALMAVLEIQNGDTRVTEEAIAARHSLSDPMNREFLCTCLADDPARR 300
DB 241 TAVVDIFSFGMCALMAVLEIQNGDTRVTEEAIAARHSLSDPMNREFLCTCLADDPARR 300
QY 301 PSASHLLEFHRVLFVYHSLKTLAAHCFIOHQYLPMPENVVEKTKAMDHLAVLAELPRPRP 360
DB 301 PSASHLLEFHRVLFVYHSLKTLAAHCFIOHQYLPMPENVVEKTKAMDHLAVLAELPRPRP 360
QY 361 PIQWRYSEVSEFMELEKFLIEDVANGIYPLMNPATRPGLPRVLADPPEEVOAKTPTDEP 420
DB 361 PIQWRYSEVSEFMELEKFLIEDVANGIYPLMNPATRPGLPRVLADPPEEVOAKTPTDEP 420
QY 421 PFSERKRYIQMOCNERSEDKARWHLTLVLVEDRLHQLTYDLLPTDSADQLASELVHY 480
DB 421 PFSERKRYIQMOCNERSEDKARWHLTLVLVEDRLHQLTYDLLPTDSADQLASELVHY 480
QY 481 GFLHEDDRMKLAFLSTFLKXRGTOA 507
DB 481 GFLHEDDRMKLAFLSTFLKXRGTOA 507

RESULT 2
ADV97865
ID ADV97865 standard; protein; 499 AA.

XX
AC ADV97865;

DT 24-MAR-2005 (first entry)

XX
DE Murine protein kinase enzyme Seq 185.

XX enzyme; protein kinase modulator; bioinformatics; diagnostic; cancer;
XX central nervous system disease; psychiatric disorder;
XX neurological disorder; neurodegenerative disease; metabolic disorder;
XX cardiovascular disease; inflammatory disorder; gene therapy; cytostatic;
XX antimigraine; analgesic; endocrine-gen.; nootropic; tranquilizer;
XX hypotensive; neuroprotective; antiparkinsonian; virocid;
XX fungicide; antibacterial; antidiabetic; anorectic; antiarteriosclerotic;
XX ophthalmological; antiinflammatory; antiallergic; antineumatic;
XX antiaesthetic; osteopathic; antipsoriatic; immunosuppressive;
XX cardiovascular-gen.; vasotropic; antiallergic; gastrointestinal-gen.;

XX
OS Mus sp.

XX
PN WO200500200-A2.
XX
XX 06-JAN-2005.
PD
XX 07-MAY-2004; 2004WO-US014421.
PF
XX 09-MAY-2003; 2003US-0469014P.
PR
XX (SUGEN) SUGEN INC.
PA
XX Caenepel S, Manning G, Charyczak G, Grigoriev I;
PI N-PSDB; ADV97751.
DR WPI; 2005-066508/07.
XX
XX
PT New isolated, enriched, or purified kinase nucleic acids and
PT polypeptides, useful for diagnosing or treating, e.g. cancers,
PT neurological and neurodegenerative diseases, cardiovascular disease, or
PT inflammatory disorders.

PS Claim 3; SEQ ID NO 185; 300bp; English.

CC This invention relates to novel isolated, enriched or purified nucleic
CC acid molecules that encode kinase polypeptides. Specifically, it refers
CC to a bioinformatics strategy used to identify mammalian members of the
CC protein and lipid kinase families. The present invention provides methods
CC for identifying a substance that modulates the activity of a kinase
CC polypeptide, as well as a method for the detection of a kinase nucleic
CC acid in a sample as a diagnostic tool for a disease or disorder.
CC Furthermore, it describes generation of a knock-out mouse whose genome is
CC disrupted by recombination at a nucleic acid sequence such that it
CC produces a phenotype, relative to the wild-type, that exhibits an absence
CC of kinase activity. The nucleic acids and polypeptides given in the
CC specification are useful for the diagnosis and treatment of cancer,
CC central or peripheral nervous system diseases, psychiatric and neurological
CC disorders, neurodegenerative diseases, metabolic disorders,
CC cardiovascular disease or inflammatory disorders. As such, they can be
CC used for gene therapy purposes and compositions exhibit cytostatic,
CC antimigraine, analgesic, endocrine-gen., nootropic, tranquilizer,
CC hypotensive, neuroprotective, antiparkinsonian, virocid,
CC fungicide, antibacterial, antidiabetic, anorectic, antiarteriosclerotic,
CC ophthalmological, antiinflammatory, antiallergic, antineumatic,
CC antiaesthetic, osteopathic, antipsoriatic, immunosuppressive, cardiant-
CC gen, vasotropic, antiallergic and gastrointestinal-gen. This polypeptide
CC is a murine protein kinase of the invention.

XX
SQ Sequence 499 AA;

Query Match 93.6%; Score 2499; DB 9; Length 499;

Best Local Similarity 93.9%; Pred. No. 5,5e-234; Indels 8; Gaps 2;

Matches 476; Conservative 11; Mismatches 12; Indels 8; Gaps 2;

QY 1 MAPEPAPRRAREREREDESEDESDESDILEESPCCGMQKRREQVNGNMPGIQSTFLAMD 60
DB 1 MAPEPAPRRAREREREDESEDESDESDILEESPCCGMQKRREQVNGNMPGIQSTFLAMD 58
QY 61 TEEGVEVWVWNEHLHFGDRKAPAAHEEKIQTFVEQVLVDHPNIVKLHKWLDTSEACAVY 120
DB 59 TEEGVEVWVWNEHLHFGDRKAPAAHEEKIQTFVEQVLVDHPNIVKLHKWLDTSEACAVY 118
QY 121 FITEVSSGSLKQPLKTKKQNHKAMNARAKMCTQIISALSFTHACSPPIIHGULTSDT 180
DB 119 FITEVSSGSLKQPLKTKKQNHKAMNARAKMCTQIISALSFTHACSPPIIHGULTSDT 178
QY 181 IFIQHNGLIKIGSVWHRIFSNALRPPTALPDDLSPIRAEREELRNLFPPPEYGEVADG 240
DB 179 IFIQHNGLIKIGSVWHRIFSNALRPPTALPDDLSPIRAEREELRNLFPPPEYGEVADG 232
QY 241 TAVVDIFSFGMCALMAVLEIQNGDTRVTEEAIAARHSLSDPMNREFLCTCLADDPARR 300
DB 233 TAVVDIFSFGMCALMAVLEIQNGDTRVTEEAIAARHSLSDPMNREFLCTCLADDPARR 292

Qy	301	PSAHSLLFRRVYLPVYSLKTLAAHCGIOYQLAPENVVEKTKAMDLHVALELPPRRP	360
Dd	293	PSAHNLLFRRLVLPVYSLKTLAAHCGIOYQLAPENVVEKTKAMDLHVALELPPRRP	352
Qy	361	PLQWYSEVSEFNEELDKLELDVANGIYPLNNFAATRLGLPRVLAAPPEEYQAKTPTPEP	420
Dd	353	PLQWYSEVSEFNEELDKLELDVANGIYPLNNFAATRLGLPRVLAAPPEEYQAKTPTPEP	412
Qy	421	FDSEFRKTYIQMOCNLERSEDKARMLTLLVLVEDRLHRLQTLTLLPTDSAQDLASLIVHY	480
Dd	413	FDSEFRKTYIQMOCNLERSEDKARMLTLLVLVEDRLHRLQTLTLLPTDSAQDLASLIVHY	472
Qy	481	GFLEHDDRMKLAAPLESTFLKRGTOA	507
Dd	473	GFLEHDDRMKLAAPLESTFLKRGTOA	499
RESULT 3			
ID	AAB65654	standard; protein; 462 AA.	
AC	AAB65654;		
XX			
DT	27-MAR-2001	(first entry)	
DE	Novel protein kinase, SEQ ID NO: 181.		
XX			
KW	Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;		
KW	immunopressiveive; cardiac; renal; antiinflammatory; antiastrmatic;		
KW	dermatological; antidiabetic; antifertility; gene therapy; vaccine;		
KW	immune disorder; cardiovascular disease; neurodegenerative disease;		
KW	cancer; autoimmune disorder; stroke; inflammatory bowel disease;		
KW	inflammatory pelvic disease; multiple sclerosis; psoriasis.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200073469-A2.		
PD	07-DEC-2000.		
XX			
PF	26-MAY-2000; 2000WO-US014842.		
XX			
PR	28-MAY-1999; 99US-0136503P.		
XX			
PA	(SUGEN-) SUGEN INC.		
PI	Plowman GD, Martinez R, Whyte D, Sudersanam S;		
DR	WPI, 2001-032161/04.		
XX			
DR	N-PSDB; AAF44681.		
XX			
PT	Nucleic acids encoding kinase polypeptides, useful for diagnosing and		
PT	treating immune-related diseases and disorders, cardiovascular disease,		
PT	neurodegenerative diseases and/or cancers.		
XX			
PS	Claim 10; Fig 1; 310pp; English.		
XX			
CC	The present sequence is a novel protein kinase. The novel protein kinases		
CC	and the nucleic acids that encode them may be used in the treatment and		
CC	diagnosis of diseases associated with inappropriate kinase expression and		
CC	such as immune-related diseases and disorders, cardiovascular disease,		
CC	neurodegenerative diseases and/or cancers. The nucleic acids and		
CC	complementary sequences may also be used as DNA probes in diagnostic		
CC	assays. The kinase polypeptides may be used as antigens in the production		
CC	of antibodies of kinase expression and activity. Anti-kinase antibodies		
CC	and kinase antagonists may also be used to down regulate kinase		
CC	expression and activity. Diseases related to kinase expression and		
CC	activity include rheumatoid arthritis, atherosclerosis, autoimmune		
CC	disorders, complications of organ transplantation, myocardial infarction,		
CC	immune disorders, cardiomyopathies, strokes, renal failure, oxidative-		
CC	stress related disorders, chronic inflammatory bowel disease, chronic		
CC	inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,		
CC	psoriasis, rhinitis, autoimmune diabetes, cancers and reproductive		

CC	disorders
XX	disorders
XX	Sequence 462 AA;
XX	Query Match 90.2%; Score 2409; DB 4; Length 462; Best Local Similarity 98.3%; Pred. No. 2,9e-225; Matches 460; Conservative 2; Mismatches 0; Indels 6; Gaps 1
QY	40 RREOVNQNMFGLOSTFLAMDTEBGEVYVMNELHFGDKAPAAHEKIQVFEQVLVDH 99
DB	1 QREKYNQNMFGLOSTFLAMDTEBGEVYVMNELHFGDKAPAAHEKIQVFEQVLVDH 60
QY	100 PNIVYLNKRWLDTSBACARVIFITEYVSSGSLKQPLKTKTKQHKMNRAMRWCTQLLS 159
DB	61 PNIVYLNKRWLDTSBACARVIFITEYVSSGSLKQPLKTKTKQHKMNRAMRWCTQLLS 120
QY	160 ALSFYHACSPPIIHGNLTSDTIFIOHNGLIKIGSVYHRIIFSVALRPPALPDDLSPIRA 219
DB	121 ALSFYHACSPPIIHGNLTSDTIFIOHNGLIKIGSVYHRIIFSVALRPPALPDDLSPIRA 174
QY	220 EREELRNHLHFPPEYGEVADGTAVDIFSGFMCALEMAVLEIQNGDTRVTEBAIARASH 279
DB	175 EREELRNHLHFPPEYGEVADGTAVDIFSGFMCALEMAVLEIQNGDTRVTEBAIARASH 234
QY	280 LSDPNMRFFIICCLARPARRPSASHLIFHVLFEVHSLKLLAAHCFIQHOTLPENYVE 339
DB	235 LSDPNMRFFIICCLARPARRPSASHLIFHVLFEVHSLKLLAAHCFIQHOTLPENYVE 294
QY	340 EKTAMDHAVALAEPRRPRPPLQMRYSSEVSFMELDKFLSDVRNGIYPLNMPAATRPGL 399
DB	295 EKTAMDHAVALAEPRRPRPPLQMRYSSEVSFMELDKFLSDVRNGIYPLNMPAATRPGL 354
QY	400 PRVLAPPEEYQAKTPTPEPDESETRKVIQMCNLESESDKAAWHLTLLVLEDRLARQ 459
DB	355 PRVLAPPEEYQAKTPTPEPDESETRKVIQMCNLESESDKAAWHLTLLVLEDRLARQ 414
QY	460 LTYDLLPTDSAQDLASELVHYGFLEHDDRMKLLAAFLSESTFLKRGTOA 507
DB	415 LTYDLLPTDSAQDLASELVHYGFLEHDDRMKLLAAFLSESTFLKRGTOA 462
RESULT 4	
ID	AD129261 standard; protein; 462 AA.
XX	AD129261;
AC	AD129261;
XX	AD129261;
XX	AD129261;
DT	22-APR-2004 (first entry)
XX	22-APR-2004 (first entry)
XX	22-APR-2004 (first entry)
DB	Human MARK3-associated protein #45.
XX	Human MARK3-associated protein #45.
KW	Human; antisense gene therapy; MARK3;
KW	MAK/microtubule affinity-regulating kinase 3; cancer;
KW	Alzheimer's disease; neurodegenerative disorder;
KW	hyperproliferative disorder; cytostatic.
XX	hyperproliferative disorder; cytostatic.
OS	Homo sapiens.
XX	Homo sapiens.
PN	US2003232771-A1.
XX	US2003232771-A1.
PD	18-DEC-2003.
XX	18-DEC-2003.
PF	17-JUN-2002; 2002US-00174319.
XX	17-JUN-2002; 2002US-00174319.
XX	17-JUN-2002; 2002US-00174319.
XX	17-JUN-2002; 2002US-00174319.
PA	(ISIS-) ISIS PHARM INC.
XX	(ISIS-) ISIS PHARM INC.
PI	Ward DT, Freiler SM, Dobie KW;
XX	Ward DT, Freiler SM, Dobie KW;
DR	WPI, 2004-052188/05.
XX	WPI, 2004-052188/05.
DR	N-PSDB; AD129379.
XX	N-PSDB; AD129379.

PT New antisense compound targeted to a nucleic acid molecule encoding
PT microtubule-affinity-regulating kinases (MARK3), useful for modulating
PT expression of MARK3 or for treating cancer or Alzheimer's disease.

XX Disclosure, SEQ ID NO 181; 233bp; English.

XX The invention relates to a compound comprising a sequence comprising 8-80
CC base pairs (bp) targeted to a nucleic acid encoding MARK3
CC (MAP/microtubule affinity-regulating kinase 3), that specifically
CC hybridises with the nucleic acid encoding MARK3 and inhibits expression
CC of MARK3, i.e. is an antisense oligonucleotide (AO). Also included are a
CC composition comprising the compound and a carrier or diluent, inhibiting
CC the expression of MARK3 in cells or tissues, treating an animal having or
CC suspected of having a disease or condition associated with MARK3 and
CC screening for an antisense compound. The antisense oligonucleotide is
CC useful for preparing a composition for treating hyperproliferative
CC disorder, particularly cancer and neurodegenerative diseases e.g.
CC Alzheimer's disease. The present sequence is a MARK3 associated protein
CC included in the figures but not mentioned anywhere else in the
CC specification.

XX Sequence 462 AA;

Query Match 90.2%; Score 2409; DB 8; Length 462;
Best Local Similarity 98.3%; Pred. No. 2,9e-225;
Matches 460; Conservative 2; Mismatches 0; Indels 6; Gaps 1;

QY 40 RREOVNQGMPGLQSTFLAMDTREGEVYVWNLHFGDRKAFANHEKIQVTEQVLVDH 99
DB 1 QREKVNQGMPLQSTFLAMDTREGEVYVWNLHFGDRKAFANHEKIQVTEQVLVDH 60
QY 100 PNVVLHKYMLDTSEACARVIFITEVSSGSLKQFLKTKKNNKAMNAAAMKRMCTQILS 159
DB 61 PNVVLHKYMLDTSEACARVIFITEVSSGSLKQFLKTKKNNKAMNAAAMKRMCTQILS 120
QY 160 ALSFLHAGSPRIHGNLTSDTFIQNGLIKIGSVWHRIJSVALRPPALPDLRSPIRA 219
DB 121 ALSFLHAGSPRIHGNLTSDTFIQNGLIKIGSVWHRIJSN-----ALPDLRSPIRA 174
QY 220 REEELRNHFFPEYGEVADGTAVDIFSGMCALENAVLEIQNGDTRVTEAIAIRAHSS 279
DB 175 REEELRNHFFPEYGEVADGTAVDIFSGMCALENAVLEIQNGDTRVTEAIAIRAHSS 234
QY 280 LSDPNMRREFILCCLARDPARPSAHSILFHRVLFVHSLKLAACFIQHOYLMENYVE 339
DB 225 LSDPNMRREFILCCLARDPARPSAHSILFHRVLFVHSLKLAACFIQHOYLMENYVE 294
QY 340 EKTAMDLHAVALAEILPRPRPPLQWRYSSEVSFMEIDKFLLEDVNRNGIYPLMNPAAATRPICL 399
DB 295 EKTAMDLHAVALAEILPRPRPPLQWRYSSEVSFMEIDKFLLEDVNRNGIYPLMNPAAATRPICL 354
QY 400 PRVLAAPPEEVOKATPTPEPDSERKVIQMCNLERSEDKARHNTLLVLERLRHQ 459
DB 355 PRVLAAPPEEVOKATPTPEPDSERKVIQMCNLERSEDKARHNTLLVLERLRHQ 414
QY 460 LTYDDLPTDSADLASLVHGFLEHDDRMKLAFLSTFLKYRGTOA 507
DB 415 LTYDDLPTDSADLASLVHGFLEHDDRMKLAFLSTFLKYRGTOA 462

RESULT 5
AAM78691
ID AAM78691 standard; protein; 357 AA.
AC AAM78691;
XX
XX 06-NOV-2001 (first entry)
DT Human protein SEQ ID NO 1353.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KM tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US004098.

XX 03-FEB-2000; 2000US-00496914.

XX 27-APR-2000; 2000US-00560875.

XX 20-JUN-2000; 2000US-00598075.

XX 19-JUL-2000; 2000US-00620325.

XX 01-SEP-2000; 2000US-00654936.

XX 15-SEP-2000; 2000US-00663861.

XX 20-OCT-2000; 2000US-00693325.

XX 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Dirmacac RT, Abundi V, Zhou P, Xu C, Cao Y;

XX Ma Y, Zhao Qa, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX Xue AJ, Yang Y, Wejhtman T, Goodrich R;

XX WPI; 2001-476283/51.

XX N-PSDB; AAK51824.

XX Claim 20; Page 3598-3599; 6221dp; English.

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit actively elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication

XX Sequence 357 AA;

Query Match 69.6%; Score 1858; DB 4; Length 357;
Best Local Similarity 97.8%; Pred. No. 9.4e-172;
Matches 355; Conservative 1; Mismatches 1; Indels 6; Gaps 1;

QY 145 MNABAMKRMCTQILSALSFHAGSPRIHGNLTSDTFIQNGLIKIGSVWHRIJSNLR 204

DB 1 MNABAMKRMCTQILSALSFHAGSPRIHGNLTSDTFIQNGLIKIGSVWHRIJSN--- 57

QY 205 PPTLAPDDLRSPITAEERBELNHLFFPEYGEVADGTAVDIFSGMCALENAVLEIQNG 264

DB 58 ---ALPDLRSPITAEERBELNHLFFPEYGEVADGTAVDIFSGMCALENAVLEIQNG 114

QY 265 DTRVTEAIAARARSHLSDPNMRREFILCCLARDPARPSAHSILFHRVLFVHSLKLAH 324

DB 115 DTRVTEAIAARARSHLSDPNMRREFILCCLARDPARPSAHSILFHRVLFVHSLKLAH 174

QY 325 CFIOHOYLMENYVEEKTAKMDLHAVALAEILPRPRPPLQWRYSSEVSFMEIDKFLLEDVANG 384

DB 175 CFIOHOYLMENYVEEKTAKMDLHAVALAEILPRPRPPLQWRYSSEVSFMEIDKFLLEDVANG 234

QY 385 IYPLMNPAAATRPICLPRVLAAPPEEVOKATPTPEPDSERKVIQMCNLERSEDKARW 444

DB 235 IYPLMNPAAATRPICLPRVLAAPPEEVOKATPTPEPDSERKVIQMCNLERSEDKARW 294

[illegible][illegible]

AB05696
ID ABB05696 standard; protein; 535 AA.
XX
AC ABB05696;
XX
DT 30-APR-2002 (first entry)
XX
DE Human brain derived protein clone fbr2_76d18.
XX
KW Human; foetal brain; foetal kidney; melanoma; testis; amygdala;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN WO200198454-A2.
XX
PD 27-DEC-2001.
XX
PF 25-APR-2001; 2001WO-IB002050.
XX
PR 25-APR-2000; 2000US-0199380P.
XX
PI (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
PI Wiemann S;
XX
DR MPI, 2002-055860/07.
XX
DR N-PSDB; ABA93734.
XX
PT Human cDNA sequences and clones derived from human fetal brain, fetal
PT kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic
PT screening and therapy.
XX
PS Claim 1; Page 216-217; 611pp; English.
XX
XX The present invention describes assemblages and computer readable media
XX comprising novel human cDNA sequences and clones derived from human
XX foetal brain, foetal kidney, melanoma, testis and amygdala cDNA
XX libraries. ABA93702 to ABA93766 represent human cDNA sequences from the
XX present invention which encode the proteins given in ABB05662 to
XX ABB05729. The human cDNA sequences and clones can be used in gene
XX therapy. The clones may be used in a variety of applications, for example
XX they may be used in profiling assays, for providing large arrays of human
XX genetic material for implementing large-scale screening strategies and
XX for treating diseases via gene therapy procedures
XX
SQ Sequence 535 AA;
XX
Query Match 57.5%; Score 1534.5; DB 5; Length 535;
Best Local Similarity 61.6%; Pred. No. 5,6e-140;
Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6;
XX
QY 17 EREDESESDILLESPCGMRKREOVNQNNPGLSTFLAMDTREGEVWNNELHFGD 76
DB 44 EEESESESESLLESPPCGMRKREOVNQNNPGLSTFLAMDTREGEVWNNELHFGD 103
QY 77 RRAFAAHEKIQTVEBQVLVDHPIVILKHTWLDTSACARVITTEYSSGSLKQFLK 136
DB 104 RKNVYLQEBKRAVAFNPLILQLEHINIVKFKHWADIKENKARVIFITEYSSGSLKQFLK 163
QY 137 KTKKNNHKNANARAKMKTQIISALSFACSPRIHGNLTGDTPIQHNGLIKTSVNH 196
DB 164 KTKKNNHKNANARAKMKTQIISALSYLHSCDPIIHGNLTGDTPIQHNGLIKTSV-- 221
QY 197 RIFSNALRPPTALPDLLRSPIRAREBELNLHFPPEYGEVAD-GTAVNIPFGMKALEM 255
DB 222 -----APDINNHVKTCREQKNLHFAPEYGEVNTTAVNIPFGMKALEM 269
QY 256 AVLEIQTNGDTR-VTEBAIARAHSLSDPNMEFLICCLARDPARRPSASHLPFRVLE 314
DB 270 AVLEIQNGESSYVPEAIISSAIQLLEDPLOREFIQCLQSPARRPRLRLFPALPE 329
QY 315 VHSVLALAHNCFIQHGYLMPENVVEKTKAMDLNVLALPR-PRRPRQWRTYSVSEFME 373

DB 330 VPSILKILAHNCFIQHGYLMPENVVEKTKAMDLNVLALPR-PRRPRQWRTYSVSEFME 389
QY 374 LDKFLIEDVNGIYPLMNFATRPGLPRLVADPPEEV-----QKAKTPEPPFSDTRK 427
DB 390 LDKFLIEDVNGIYPLTAF-----GLPRPQOQOQEEVTSPPVVPVKRTPEPBAVETR 443
QY 428 VIQMCNLEERSDKARMTLLVLVEDRLHROLTVDLPTDANQDLASLHYKGFLEHD 487
DB 444 VVLMQCNLESVEBGVGHLLTLKEDKILNRHSCDLMENENIPELAAELVOLGFISEAD 503
QY 488 RMLKAAFLSESTFLKY 502
DB 504 QSRUTSLEETLANKF 518
XX
RESULT 10
ADD47447
ID ADD47447 standard; protein; 535 AA.
XX
AC ADD47447;
XX
DT 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX
DE Human Protein NP_037524, SEQ ID NO 13142.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNJ; Chung.
XX
OS Homo sapiens.
OS Unidentified.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHU-) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR MPI, 2003-268312/26.
DR GENBANK; NP_037524.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Example 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the

Qy 374 LDKFEDVANGIYPLNNFATRLPLGLPVLAAPPPEV-----QKAKTPPEPDSSETRK 427
Db 390 LDKFEDVANGIYPLTAFA-----GLPRPOQPOQEEVTSVPVPSVKTPPEPAVEYTRK 443
Qy 428 VIOMQCNLERSDEKARWHLTLVLVEDRLHRLQTLTYLLPTDSAQDLASLWVHGFLEDD 487
Db 444 VVLMQCNISVEGVGHHTLLKLEDKLNRLHSLCOLMPENIPELAELVQLGFTSEAD 503
Qy 488 RMKLAFFLESTFLKY 502
Db 504 QSRLLTSLLEETLNKF 518
RESULT 12
AD129264
ID AD129264 standard; protein; 535 AA.
AC AD129264;
XX
XX 22-APR-2004 (first entry)
DT
XX
XX Human MARK3-associated protein #47.
DE
XX
XX Human; antisense gene therapy; MARK3;
KW MAP/microtubule affinity-regulating kinase 3; cancer;
KW Alzheimer's disease; neurodegenerative disorder;
KW hyperproliferative disorder; cytostatic.
XX
OS Homo sapiens.
XX
XX US2003232771-A1.
XX
XX 18-DEC-2003.
XX
XX 17-JUN-2002; 2002US-00174319.
XX
XX 17-JUN-2002; 2002US-00174319.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Ward DT, Freiler SM, Dobie KW;
PI
XX
XX WPI; 2004-052188/05.
DR
XX
XX N-PSDB; AD129382.
XX
XX
XX New antisense compound targeted to a nucleic acid molecule encoding
PT microtubule-affinity-regulating kinases (MARK3), useful for modulating
PT expression of MARK3 or for treating cancer or Alzheimer's disease.
XX
XX
XX Disclosure; SEQ ID NO 184; 233pp; English.
XX
XX
XX The invention relates to a compound comprising a sequence comprising 8-80
CC base pairs (bp) targeted to a nucleic acid encoding MARK3
CC (MAP/microtubule affinity-regulating kinase 3), that specifically
CC hybridizes with the nucleic acid encoding MARK3 and inhibits expression
CC of MARK3, i.e. is an antisense oligonucleotide (AO). Also included are a
CC composition comprising the compound and a carrier or diluent, inhibiting
CC the expression of MARK3 in cells or tissues, treating an animal having or
CC suspected of having a disease or condition associated with MARK3 and
CC screening for an antisense compound. The antisense oligonucleotide is
CC useful for preparing a composition for treating hyperproliferative
CC disorder, particularly cancer and neurodegenerative diseases e.g.
CC Alzheimer's disease. The present sequence is a MARK3 associated protein
CC included in the figures but not mentioned anywhere else in the
CC specification.
XX
XX
XX Sequence 535 AA;
SQ
Query Match 57.5%; Score 1534.5; DB 8; Length 535;
Best Local Similarity 61.6%; Pred. No. 5,6e-140;
Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6;
Qy 17 BREDESEDSILBESPCGRWQKRGROVNVQGNMPLGQSTFLAMDTEGVEVVMNELHFGD 76

Db 44 EEESEDESEILBESPCGRWQKRGROVNVQGNMPLGQSTFLAMDTEGVEVVMNELHFGD 103
Qy 77 RKAFAAHEEKIOTVEQOLVVDHPNIVKLHKTWLDTSEACANVIFITEYVSSGSLKQFLK 136
Db 104 RKNYKLOEERKAVAFONLILQSHLNIIVKFKHYWADIKENKARVIFITEYVSSGSLKQFLK 163
Qy 137 KTKKNNKANNARAKMRCQOILSALSFACSPPIIHGNTLSDTTFIQHNGLIKIGSVH 196
Db 164 KTKKNNKNNERAKMRCQOILSALSYLHSCDPIIHGNTLSDTTFIQHNGLIKIGSVH 221
Qy 197 RIFSNALRPPTALPDRLRSPRAREELNHLFPPEYGEVAD-GTAVDIFSGMCALPM 255
Db 222 -----APDTINNHVKTCREQNLHFPABEYEVNVTVAVDIYFSGCALPM 269
Qy 256 AVLEIQTNQDTR-VTEBAIARARHSISDNMEFFLCCILARDPARPSASHLLFHVLEP 314
Db 270 AVLEIQNGSESSYVPEAIISSAIIQLLEDPLQREFFIQKIQSEPARPFTARELLFHPALPE 329
Qy 315 VHSKLILAAHCFIQOYLMPENVVEKTKAMDHLAVLALPR-PRRPFIQMKYSEVSM 373
Db 330 VPSLKLILAAHCFIQOYHMIPEVALBEITQMDTSAVLAEIPAGPGRPEVQTLYSQSPALPE 389
Qy 374 LDKFEDVANGIYPLNNFATRLPLGLPVLAAPPPEV-----QKAKTPPEPDSSETRK 427
Db 390 LDKFEDVANGIYPLTAFA-----GLPRPOQPOQEEVTSVPVPSVKTPPEPAVEYTRK 443
Qy 428 VIOMQCNLERSDEKARWHLTLVLVEDRLHRLQTLTYLLPTDSAQDLASLWVHGFLEDD 487
Db 444 VVLMQCNISVEGVGHHTLLKLEDKLNRLHSLCOLMPENIPELAELVQLGFTSEAD 503
Qy 488 RMKLAFFLESTFLKY 502
Db 504 QSRLLTSLLEETLNKF 518
RESULT 13
AD058699
ID AD058699 standard; protein; 535 AA.
AC AD058699;
XX
XX
XX 15-JUL-2004 (first entry)
DT
XX
XX Human regulatory molecule HRM-20.
DE
XX
XX cytostatic; immunomodulator; agonist; antagonist; gene therapy;
KW human regulatory molecule; HRM; disease development; cell proliferation;
KW immune response; cancer.
XX
XX
XX Homo sapiens.
OS
XX
XX US2002058264-A1.
XX
XX 16-MAY-2002.
XX
XX 26-SEP-2001; 2001US-00840787.
XX
XX 23-SEP-1997; 97US-00933750.
XX
XX 20-JAN-1999; 99US-00234613.
XX
XX 03-MAR-2000; 2000US-00518865.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX
XX Lal P, Hillman JL, Bandman O, Shah P, Au-Young J, Yue H;
PI Guegler KJ, Corley NC;
XX
XX WPI; 2004-459763/43.
DR
XX
XX N-PSDB; AD058748.
XX
XX
XX New human regulatory molecules, useful in the diagnosis and treatment of
PT cancer and immune disorders.
XX

PS Claim 1; SEQ ID NO 20; 116pp; English.
 XX The invention describes human regulatory molecules (HRM) (I) selected
 CC from a group comprising the fully defined amino acid sequences of SEQ ID
 CC NOs: 1-49. Also described are: an isolated polynucleotide (II) comprising
 CC a nucleic acid sequence encoding (i) or the complement of the
 CC polynucleotide (SEQ ID NOs:50-98); a composition comprising (II) and a
 CC reporter molecule; an expression vector containing (II); a host cell
 CC containing the vector; detecting (M1) expression of a nucleic acid in a
 CC sample; screening (M2) a plurality of molecules to identify a ligand;
 CC diagnosing (M3) a disease associated with gene expression in a sample
 CC containing nucleic acids; a composition comprising (I) and a
 CC pharmaceutical carrier or a labeling moiety; screening (M4) a plurality
 CC of molecules to identify a ligand; preparation and purification of
 CC antibodies; an antibody which specifically binds to (i); and detecting
 CC protein expression in a sample. The new human regulatory protein
 CC molecules which are expressed during disease development and the
 CC polynucleotides which encode them satisfies a need in the art by
 CC providing compositions which are useful in the diagnosis and treatment of
 CC diseases associated with cell proliferation, particularly immune
 CC responses and cancer. This is the amino acid sequence of a human
 CC regulatory molecule.
 XX
 XX Sequence 535 AA;

Query Match 57.5%; Score 1534.5; DB 8; Length 535;
 Best Local Similarity 61.6%; Pred. No. 5.6e-140;
 Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6;

17 EREDESESDILSESPCGRMQKREOVNQNMPGLQSTFLAMDTEEGEVVWNNELHFGD 76
 44 EEESESESEILSESPCGRMQKREOVNQNMPGLQSTFLAMDTEEGEVVWNNELHFGD 103
 77 RKAFAHEKIQTFEQLVLDHPNIVKHKWLDSEACAVITFEVSSGSLKQFLK 136
 104 RKNYKLGEBKVAEPNLIQLEHLNIVKHKWLDSEACAVITFEVSSGSLKQFLK 163
 137 KTKKHKAMNARAKRWCTQILSALSFHACSPRIHGNLTSDTIFIOHNGLIKIGSVH 196
 164 KTKKHKMTNEKAKRWCTQILSALSYLHSCDPIIHGNLTCDTIFIOHNGLIKIGSV-- 221
 197 RIFSNALRPPTALPDLSRPAERERELNHPPEVEGVAD-GTAVDIFSGKCALEM 255
 222 -----APDTINNHVKTCEBQKHLHFPAEYEVNNTTAVDIYSFGMCALAM 269
 256 AVLEIOTNGDTR-VTBEAIAARAHSLSDPMREFFICCLARPAPRPSHSLFRVLF 314
 270 AVLEIOTNGESSYVPEALISSAIQLLEDPLOREFFIOKLOSEPAPRPTARELLFFPALFE 329
 315 VHSKLTLAAHCFIOHQLMPENVEEKTAMDHLAVIAELPR-PRRPLQMYSEVSEFMS 373
 330 VPSLKLTLAAHCFIOHQMHPENALBEITQNMOTSAVLAIPAGPQREPTQISQSPALF 389
 374 LDKFLEADVNGIYPLMNPATRPGLPRVLAPRPEEV-----QKATPTPEPPESETRK 427
 390 LDKFLEADVNGIYPLTAF-----GLPRPOQEQEVEVTPVVPSPKPTPEPALAEVETRK 443
 428 VTQMCNTERSEEDKARWHLTLVLVEDRLHROTLYDILPTDQAOLASVHYGFLHEDD 487
 444 VVLMQCNTERSEEGVGHRLTLKLKEDKLNRLSCDLMNENIPELAAELVOLGFISEAD 503
 488 RMLKLAFLSESTFLKY 502
 504 QSRITSLSLEETLNKF 518

RESULT 14
 ID AEA61758 standard; protein; 535 AA.
 XX AEA61758;
 AC AEA61758;
 XX
 DT 11-AUG-2005 (first entry)

XX Human nuclear receptor binding protein (NRBP).
 DE
 XX Nuclear receptor binding protein; GTPase modulator; cancer; neoplasm;
 KM cytosolic; gene therapy; antibody therapy; drug screening; diagnosis.
 XX
 OS Homo sapiens.
 PN W0205052130-A2.
 XX
 XX 09-JUN-2005.
 PD
 PF 23-NOV-2004; 2004WO-US039708.
 PF
 PR 24-NOV-2003; 2003US-0524541P.
 XX
 PA (EXEL-) EXBLIXIS INC.
 PI Kadyk LC, Francis GR, Lickteig K;
 XX
 DR WPI: 2005-417982/42.
 DR N-PESDB; AEA61756.
 DR REFSBQ; NP_037524.
 XX
 PT Identifying a candidate RAC pathway modulating agent using a nuclear
 PT receptor binding protein (NRBP) polypeptide or nucleic acid, useful for
 PT diagnosing or treating cancer.
 XX
 PS Example 2; SEQ ID NO 3; 51pp; English.

The present sequence is the protein sequence of human nuclear receptor binding protein (NRBP), a ubiquitously expressed, multidomain adapter protein. The inventors have discovered genes that modify the RAC GTPase pathway in *Caenorhabditis elegans* and identified their human orthologs as CC NRBP. The invention provides methods for using these RAC modifier genes CC and polypeptides to identify NRBP-modulating agents that are candidate CC therapeutic agents that can be used in the treatment of disorders CC associated with defective or impaired RAC function and/or NRBP function. CC Preferred NRBP-modulating agents specifically bind to NRBP polypeptides CC and restore RAC function. Other preferred NRBP-modulating agents are CC nucleic acid modulators such as antisense oligomers and RNAi that repress CC NRBP gene expression or product activity. NRBP modulating agents may be CC evaluated by *in vitro* or *in vivo* assay for molecular interaction with an CC NRBP polypeptide or nucleic acid. Agents that produce a change in the CC activity of these assay system relative to controls are identified as CC candidate RAC modulating agents. The assay system may be cell-based or CC cell-free. NRBP modulating agents include NRBP-related proteins (e.g. CC dominant negative mutants and biotherapeutics), NRBP-specific antibodies, CC NRBP-specific antisense oligomers and other nucleic acid modulators, and CC chemical agents that specifically bind to or interact with NRBP or CC apoptosis assay, a cell proliferation assay, an angiogenesis assay or a CC hypoxic induction assay. A method for diagnosing a disease, especially CC cancer, in a patient uses a probe for NRBP expression.
 XX
 XX Sequence 535 AA;

Query Match 57.5%; Score 1534.5; DB 9; Length 535;
 Best Local Similarity 61.6%; Pred. No. 5.6e-140;
 Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6;

17 EREDESESDILSESPCGRMQKREOVNQNMPGLQSTFLAMDTEEGEVVWNNELHFGD 76
 44 EEESESESEILSESPCGRMQKREOVNQNMPGLQSTFLAMDTEEGEVVWNNELHFGD 103
 77 RKAFAHEKIQTFEQLVLDHPNIVKHKWLDSEACAVITFEVSSGSLKQFLK 136
 104 RKNYKLGEBKVAEPNLIQLEHLNIVKHKWLDSEACAVITFEVSSGSLKQFLK 163
 137 KTKKHKAMNARAKRWCTQILSALSFHACSPRIHGNLTSDTIFIOHNGLIKIGSVH 196
 164 KTKKHKMTNEKAKRWCTQILSALSYLHSCDPIIHGNLTCDTIFIOHNGLIKIGSV-- 221

PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-581633/65.
 DR N-PSDB; ABR43561.
 XX
 PT New isolated nucleic acid encoding a protein for diagnosing, preventing,
 PT creating or ameliorating medical conditions and used as food additives or
 PT preervatives.
 PS
 PS Claim 9; SEQ ID NO 749; 837pp; English.

CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
 CC acute kidney failure and blood related disorders e.g. myocardial
 CC infarction. The polypeptides can also be used to aid wound healing and
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
 CC maintain organs before transplantation, for supporting cell culture of
 CC primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content, lipid, protein,
 Query Match 57.5%; Score 1534.5; DB 4; Length 581;
 Beef Local Similarity 61.6%; Pred. No. 6.4e-140; Indels 29; Gaps 6;
 Matches 305; Conservative 59; Mismatches 102;
 QY 17 EREDESEDESDILESPCGRQKREQVNOQNMPLQSTFLAMDTEGSEVVMNLAHGD 76
 Db 90 EEESESESESELESPCGRQKREQVNOQNMPLQSTFLAMDTEGSEVVMNLAHGD 149
 QY 77 RKAPAAHEKIQTFEQLVLDHPNIVKLAHKYMDTSEACARVIFITVYSSGSIKQPLK 136
 Db 150 RKNYQLQGEKRVAVPDNLQLEHLNIVKFKHYWADIKENKARVIFITVYSSGSIKQPLK 209
 QY 137 KTKKHKAMNARAWKRWCTQILSLSPHACSPRIHGNLTSDPTFIQNGLIKGSVNH 196
 Db 210 KTKKHKMTNEMKAWRCWCTQILSLSLYHSCDPTIHGNLTCDTIFIQNGLIKGSV-- 267
 QY 197 RIFSNALRPPTALPDLSPIRAERERELRNHFPPPEYGEVAD-GTAVDIFSGMCALEM 255
 Db 268 -----ADPTINNHYKTCREBQKNIHFAPREGEVTNVTYTAIDIVSFGCALEM 315
 QY 256 AVLEIQINGDTR-VTEEAIDARRSLSDPNRBEFLCCLADDPARPSAHSILFHRVLE 314
 Db 316 AVLEIQNGESSYVPOEAISSAIOLEBDPLQREFIOKLOSEPAPRRPARELLFHPALFE 375
 QY 315 VHSKLTLAHCFOIQVMPENNVVEKTKAMDLHVALELR-PRRPLQWRYSVSME 373
 Db 376 VPSLKLALAHCTVGHQHPENALBEITXNDTSVLALEIPAGPREVQTIYSOSPALD 435
 QY 374 LDKFLEDVRNGIYPLMNFATRPGLPVLAPPREY-----QKAKTPTEPPFSETRK 427
 Db 436 LDKFLEDVRNGIYPLTAF-----GLPRPQPGQGESEVSPVPPSVKPTPEPAVAVERK 489
 QY 428 VYMOQCNLERSEBDKARWHLTLVLVEDRLHQLTYDILPTPSAODLASELVHYGFLHDD 487
 Db 490 VLMQCNIESYVEGVKHLTLTLKLEDKLNHLSCLMPNBNIPETLAELVOLGFISSAD 549
 QY 488 RMKLAFLFESTFLKY 502
 Db 550 QSRLTSLLEETLNKF 564
 Search completed: January 12, 2006, 11:20:31
 Job time : 83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 13, 2006, 21:08:50 ; Search time 6250 Seconds
(without alignments)
4611.141 Million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670
Sequence: 1 MAAPEAPAPRARERERERED.....RMKLAFLSTLTKRGTA 507

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+gch.model -DEV=xip
-Q=/cgn2.1/USPTO_epool_p/US10618941/runat_12012006_070204_14268/app_query.fasta_1.647
-DB=GenBdb1 -QFMT=faststep -SUFFIX=erge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=45
-UNIT8=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=EXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10618941@CGN_1_1_7415@runat_12012006_070204_14268 -NCPU=6 -ICPU=3
-NO_WMAP -LARGECQUERY -NEG SCORES=0 -WAIT -DSPELCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenBdb1:.*
1: gb_ba:.*
2: gb_in:.*
3: gb_env:.*
4: gb_ov:.*
5: gb_ov:.*
6: gb_ov:.*
7: gb_ov:.*
8: gb_ov:.*
9: gb_ov:.*
10: gb_ov:.*
11: gb_ov:.*
12: gb_ov:.*
13: gb_ov:.*
14: gb_ov:.*
15: gb_ov:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2612	97.8	4048	6	AR448352 Sequence
2	2575	96.4	3892	8	BC071605 Homo sapi
3	2517.5	94.3	1548	6	CQ729503 Sequence

4	2409	90.2	3304	6	AX056416	AX056416 Sequence
5	2368	88.7	3538	8	HSW601974	AL117662 Homo sapi
6	1617	60.6	981	6	AR310264	AR310264 Sequence
7	1617	60.6	981	6	AR640307	AR640307 Sequence
8	1612	60.4	92291	8	AC105049	AC105049 Homo sapi
9	1563.5	58.6	174949	14	AC116487	AC116487 Homo sapi
10	1563.5	58.6	202844	9	AC110211	AC110211 Mus muscu
11	1563.5	58.6	213817	14	AC051515	AC051515 Mus muscu
12	1555.5	58.3	194709	14	AC132640	AC132640 Rattus no
13	1555.5	58.3	248568	14	AC115184	AC115184 Rattus no
14	1541	57.7	3046	5	BC074684	BC074684 Xenopus t
15	1537.5	57.6	2128	9	BC004756	BC004756 Mus muscu
16	1537.5	57.6	2148	9	BC018463	BC018463 Mus muscu
17	1537.5	57.6	2152	9	AF302138	AF302138 Mus muscu
18	1536	57.5	2461	9	BC011468	BC011468 Mus muscu
19	1534.5	57.5	1608	11	AY335620	AY335620 Synthetic
20	1534.5	57.5	2029	6	AR116039	AR116039 Sequence
21	1534.5	57.5	2101	8	AB168327	AB168327 Macaca fa
22	1534.5	57.5	2141	6	BD156896	BD156896 Primer fo
23	1534.5	57.5	2141	6	AX877871	AX877871 Sequence
24	1534.5	57.5	2141	8	AK001946	AK001946 Homo sapi
25	1534.5	57.5	2163	6	BD186292	BD186292 STR6 act
26	1534.5	57.5	2163	6	CQ718639	CQ718639 Sequence
27	1534.5	57.5	2163	6	AX056419	AX056419 Sequence
28	1534.5	57.5	2163	6	AF113249	AF113249 Homo sapi
29	1534.5	57.5	2194	8	AB169835	AB169835 Macaca fa
30	1534.5	57.5	2197	8	BC001221	BC001221 Homo sapi
31	1534.5	57.5	2883	8	AK122664	AK122664 Homo sapi
32	1529.5	57.3	2159	6	BD156223	BD156223 Primer fo
33	1529.5	57.3	2159	6	AX876616	AX876616 Sequence
34	1529.5	57.3	2159	6	BD012203	BD012203 Novel gen
35	1529.5	57.3	2159	6	AK027538	AK027538 Homo sapi
36	1528	57.2	2537	9	BC012437	BC012437 Mus muscu
37	1527.5	57.2	2131	8	AK223136	AK223136 Homo sapi
38	1522.5	57.0	2178	6	BD186348	BD186348 STR6 act
39	1519	56.9	2810	5	AJ721116	AJ721116 Gallus ga
40	1514	56.7	2870	5	BC068805	BC068805 Xenopus l
41	1508.5	56.5	2241	6	AR253958	AR253958 Sequence
42	1465	54.9	232234	14	AC108634	AC108634 Rattus no
43	1394	52.2	157461	14	AC149780	AC149780 Bos tauru
44	1360.5	51.0	2228	6	AX056417	AX056417 Sequence
45	1346	50.4	777	8	CR457350	CR457350 Homo sapi

ALIGNMENTS

RESULT 1
AR448352
LOCUS AR448352 4048 bp DNA
DEFINITION Sequence 1053 from patent US 6673549.
ACCESSION AR448352
VERSION AR448352.1 GI:42676676
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4048)
Furness,L.M. and Buchbinder,J.L.
TITLE Genes expressed in C3A liver cell cultures treated with steroids
JOURNAL Patent: US 6673549-A 1053 06-JAN-2004;
Incyte Corporation; Palo Alto, CA
FEATURES
source
1..4048
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1,16e-213 Length: 4048
Score: 2612.00 Matches: 501
Percent Similarity: 98.62% Conservative: 0
Best Local Similarity: 98.62% Mismatches: 0
Query Match: 97.85% Indels: 7

DB:	6	Gap:	1
US-10-618-941-99 (1-507) x AD448352 (1-4048)			
OY	1	MetalAaLaProGluProAlaProAlaArgAlaArgGluAaArgGluAaArgGluAaArg	20
Db	282	ATGGGGCCCCCGAGCGCGCGCGCGAGCGCGCGCGGAAACGGAGCGCGAGCGCGAGAC	341
OY	21	GluSerGluAaArgGluSerAaPrlLeuGluGluSerProCysGlyAaGTPGlnlyAaArg	40
Db	342	GAGACCGAGGACGAGAGCGACATCCCTCGAGAGAAAGCCGTTGTGCTGGCAAAAGCGA	401
OY	41	ArgGluGlnValAaArgGlnlyAaMetProGluylLeuGlnSerThrPheLeuAaMetAaP	60
Db	402	CGGAGCGAGTAAACCAAGGAAACAGACGAGCGCTTCAAGACCTTCTTACGCAATGAGAC	461
OY	61	ThrGluGluGluValAlGluValValTTPAaArgGluLeuHisAaPheGlyAaAaPhe	80
Db	462	ACGAGAGAGGGGGTACAGGTGGTGTGGAACGAGCTCATTCCGAGAACAGAAAGGCTTTC	521
OY	81	AlaAaAaHsGluGluLyAaHsGlnThrValPheGluGlnLeuValLeuAaAaPrlAaPro	100
Db	522	GGGGCGCAGACGAGAGAAAGATCCAGACCGATGTTGACAGAGTGGTGGTGAGACACCG	581
OY	101	AaHsLeuValylLeuHisAlySerTyrTrpLeuAaPrlHisSerGluAaLyAaAaArgValIle	120
Db	582	AACATCGTGAAGTTCGACAAAGTACTCGCTGGATACCTCGAGGCGCTGGCGAGAGTTCATC	641
OY	121	PheIleThrGluTyrValSerSerGlySerLeuLyGlnPheLeuLyAaThrLyAaLyAa	140
Db	642	TTTATATCAAGAATACGTGTCAATCAGGACGCTTACAGCAATTCTTCAAAAAGCAAGAAAG	701
OY	141	AaHsHsLyAaAaMetAa	160
Db	702	AACCAACAAGGCATGAAACCCCGGGCTGGAAGCGCTGGAGCAACGAGATCTGTCTGGCG	761
OY	161	LeuSerPheLeuHisAaAlyCysSerProProIleIleHisGlyAaHsLeuThrSerAaPrlHis	180
Db	762	CTCAGCTTCTCTCAACGCTCGACGCGCCCAATCATCACGGGAACCTGACACACGACACACC	821
OY	181	IlePheIleGlnHisAaArgGlyLeuIleLyAaHsGlySerValTTPHisAaArgIlePheSer	200
Db	822	ATCTTCATTACAGACACACGCGCTCATCAAGATGGCTCCGTGTGGACCGAATCTTCTCC	881
OY	201	AaHsAaAaLeuAaArgProProThrAaLeuProAaAaPrlAaAaSerProIleAaArgAlaGlu	220
Db	882	AAT-----GACATTCAGATGATCTCCGAAGCCCAATCCGCGGTGAG	923
OY	221	ArgGluGluLeuAaArgAaHsLeuHisAaPheProProGluTyrGlyGluValAaAaArgLy	240
Db	924	CGAGAGGAATCTCGGAACCTGCACTTCTTCCCCCAGAGTATGAGAAGGTGGCGCATGGGG	983
OY	241	ThrAlaValAaPrlPheSerPheGlyMetCysAlaLeuGlnMetAaValaLeuGlnIle	260
Db	984	ACCGCTGTGGACATCTTCTCTTGGAGTGTGGCTGGAGATGGCTGTACTGAAATTC	104
OY	261	GlnThrAaArgLyAaAaPrlAaArgValThrGluGluAlaIleAaAaArgAlaAaArgHisSerLeu	280
Db	1044	CAGACCAAAAGGGGACACCCGGGTCAAGAGAGAGGCCATTCTGCGCCACAGCACTGGCTG	110
OY	281	SerAaPrlAaAaMetAaArgGluPheIleLeuCysCysLeuAlaAaAaPrlAaAaArgAa	300
Db	1104	AGTGACCCCAACATCGGGAGTTCATCTTGTCTGGCTGGCCCGGACCTTGGCCCGCGG	116
OY	301	ProSerAlaHisSerLeuLeuPrlHisAaArgValaLeuPheGluValHisSerLeuLyAa	320
Db	1164	CCCTGTGCCACACGCTCTCTTCCACCGGGTCTTTGAGAGTGAATCTGCTGAAGCTC	122
OY	321	LeuAlaAaAaHisCysAaPheIleGlnHisGlnTyrLeuMetProGluAaAaValaIleGluGlu	340
Db	1224	CTGGCAGGCCCATCTGCTTATCCAGACACACAGTACTATGCTGTAGAAATGTGTGTGAGAGAG	128
OY	341	LyAaThrLyAaAaMetAaPrlLeuHisAaValaLeuAlaGluLeuAaPrlAaArgProAaArgAaPro	360

Dd	1284	AAGACCAAGCCATGAGACTTGACGGGTCTTGGCGAGCTTCCCCGGC	1344
Oy	361	ProLeuGIINrPaYrYrSerGIuValSerPheMetGIueAspLyvPheLeuGIuaap	380
Dd	1344	CCGGTCGAGTGGCGGGTACTCGGAAGTCTCTTCATGGAGCTGAGCAAAATTCTCGGAGANT	1403
Oy	381	ValArGaNGIlylIeYrProLeuMeCaenPheAlalathrArGPProLeuGIyleuPro	400
Dd	1404	GTCAGGAATGAATCTACCCTGATGTAACTTTGGAGCACCTGCACCCCTGGGGCTGCC	1463
Oy	401	ArgValleuAlAProPProGIuGlValGIuNyvalAlayThr-ProThrProGIuPr	420
Dd	1464	CGTGGCTGGCCCCAACCCCGAGAGGTCCAAGAAGCAAAGACCCCGAGCCAGAGCC	1522
Oy	420	oPheAspSerGIuThrArGLysValIlEgImetGIInCySaenLeuGIuaYrSerGIuas	440
Dd	1524	CTTTGACTCTGAGACCAAGAAAGTCTCATGATGCAAGTGCACCTGGAGAGAGAGAGA	1583
Oy	440	pLYvalAArgTrpHisLeuThrLeuLeuValleuGIuAspArGLueHIArGLule	460
Dd	1584	CAAAGCGCGCTGGCATCTCACTGCTCTTGCTGGTGGAAACCGGCTGCACCGGACAGCT	1644
Oy	460	uThrTyrrAspLeuLeuPProThraAspSerAlIGLAspLeuAlasergIuleValHisTy	480
Dd	1644	GACCTACCACTGCTCCCACAAGGACGAGCCACAGACCTCGCTCGAGGCTGGACCTA	1703
Oy	480	rGLYPheLeuHISGLuAspAParGMetLysLeuAlAlaPheLeuGIuSerThrPheLe	500
Dd	1704	TGGCTCTCTCCACAGAGACGACCGGATGAACCTGGCCGCTTCTCGAGAGACACTTCT	1763
Oy	500	uLYeTYrArGLYThrGIuAla	507
Dd	1764	CAAGTACCCTGGAGACCCAGGCC	1785
RESUT_2			
BC071605			
LOCUS	BC071605	3892 bp	mRNA linear PRI 25-JUN-2004
DEFINITION	Homo sapiens hypothetical protein LOC340371, mRNA (cDNA clone IMAGE:4375917), partial cds.		
ACCESSION	BC071605		
VERSION	BC071605.1	GI:48735259	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE			
AUTHORS	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo 1 (bases 1 to 3892)		
	Strausberg,R.L., Collins,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heide,F., Ditchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stempleton,M., Soares,M.B., Bonaldi,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Udell,T.B., Tothbiakyl,S., Carinini,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mollahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulik,S.W., Villalón,D.K., Wuzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Faley,J., Helton,E., Keltman,M., Madan,A., Young,A.C., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bonfard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schnut,J., Myers,R.M., Butterfield,V.S., Krzywinski,M.I., Skalska,U., Smalins,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 2 (bases 1 to 3892) Strausberg,R.		

QY 439 GIUAAPLYAALAAAGTTPHISLEUThRIELeuLeuValLeuGIUAAPRGLUeunHISArg 458
 DB 1600 GAGGACAGAGCCGCGGATCTCACTCTGCTTGGTCTGGAAGACCGGCTGACCGG 1659
 QY 459 GInLeuThrTYRAspLeuLeuProThraSPSerAAGInAspLeuAlaSerGIUeunVal 478
 DB 1660 CAGCTGACCTACGACTGCTCCCAAGCGACGCGCCAGAGACTGCGCTCGAGCTCGTG 1719
 QY 479 HIRTYRGLYPheLeuHISGLIUAAPRAspArgMeLYISLeuAlaAlaPheLeuGIUeunThr 498
 DB 1720 CACTATGGCTCTCTCCACAGAGACGACCGGATGAAGCTGCGCTCTCTGAGAGACC 1779
 QY 499 PheLeuLYETRYArgGLIYThrglnAla 507
 DB 1780 TTCCTCAAGTACCGTGGAGCCAGGCC 1806
 RESULT 3
 LOCUS CQ729503 1548 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 15437 from Patent WO02068579.
 ACCESSION CQ729503
 VERSION CQ729503.1 GI:42300828
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE
 AUTHORS
 TITLE
 1
 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 Kite, such as nucleic acid arrays, comprising a majority of
 humeunons or transcripts, for detecting expression and other uses
 thereof
 Patent: WO 02068579-A 15437 06-SEP-2002;
 PE Corporation (NY) (US)
 JOURNAL
 FEATURES
 source
 1. 1548
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,286-206 Length: 1548
 Score: 2517.50 Matches: 489
 Percent Similarity: 94.22% Conservation: 0
 Best Local Similarity: 94.22% Mismatches: 1
 Query Match: 94.29% Indels: 29
 DB: 6 Gaps: 4
 US-10-618-941-99 (1-507) x CQ729503 (1-1548)

QY 112 ThRSerGIUAALCYAALAAAGValIlePheIleThrgIuTYValISerSerGIYSerLeu 131
 DB 307 ACCTTGAGGCTCTGGGAGGCTCATCTTCACTCAAGAGTACGTCTCATGAGGAGCTTC 366
 QY 132 LYSGLNpheLeuLYETHTYRLeYLYeunHISLYAALaMeTAsnAlaArgAlaTrpLYs 151
 DB 367 AAGCAATTCCTCAAAAAGACCAAGAGAACCAAGGCGCATGAAGCGCCGGGCTGGAAG 426
 QY 152 ArgTrpCYThrgInIleLeuSerAlaLeuSerPheLeuHISAlaCYSerProTrIle 171
 DB 427 CGTGGTGCAGCAGATCTGTCTGCGTCAAGCTTCTGCACGCTTGAGAGCCCTCCCAATC 486
 QY 172 ILeHISGLYAAnLeuThrSerAspThrIlePheIleGlnHISaenGIYLeuIleYsIle 191
 DB 487 ATCCAGGGGAACCTGACCGAGGACACCATCTTCACTTCAAGCACAGGCTCATCAAGATC 546
 QY 192 GLYSer-----ValTrpHISArgIlePheSerAsnAlaLeuArg 204
 DB 547 GGCCTCCGGTGTGGCGGGGAGGCTGGTGTGGCACCGAATCTTCCAAAT----- 597
 QY 205 ProProThraAlaLeuProAspAspLeuArgSerProIleArgAlaGLIUAArgGIUeun 224
 DB 598 -----GCACCTTCAGATGATCTCCGAAGCCCCCATCCGCGCTGAGCGAGAGACTT 648
 QY 225 ArgAsnLeuHISpHePheProGIUTrGLYGLIuValAlaAspGLYThraAlaValAsp 244
 DB 649 CGGAACCTGCACCTTCTTCCCCCAGAGTATGAGAGGTGGCGGATGGAGCCGCTTGAGAC 708
 QY 245 ILePheSerPheGIYMeTCYAlaLeuGLIuMeAlaValLeuGIuIleGlnThraenGIY 264
 DB 709 ATCTTCTCTTGGGATGTGTGCGCTGGAGATGGCTGTACTGAAATCCAGACCAATGGG 768
 QY 265 AspThraArgValIThrgIuGLIuAlaIleAlaArgAlaArgHISerLeuSerAspProAsn 284
 DB 769 GACACCCGGGTCAAGAGAGAGCCATGTCTGCGCGAGGCACTGCTGTAAGTGAACCCCAAC 828
 QY 285 MecArgGIUPheIleLeuCYeCYeLeuAlaArgAspProAlaArgArgProSerAlaHIS 304
 DB 829 ATGCGGAGTTCATCTTGTCTGTGCTGGCCCGGAGACCTGGCCCGGCTCTTCCAC 888
 QY 305 SerLeuLeuPheHISArgValIleuPheGIUValHISerLeuLYeLeuAlaAlaHIS 324
 DB 889 AGCTCCTCTTCCACCGGCTGCTCTTGAAGTGCACCTGTAAGCTCTGGCAAGCCAC 948
 QY 325 CYPheIleGlnHISGLINTY-----Leu 332
 DB 949 TGCCTTCACTCAAGCACCACTGTGACACAGGCGCGCGGCTGCTGGGAGACCTC 1008
 QY 333 MecProGIUAAnValIaIGIUGIuLYETHTYRValaMeTAspLeuHISAlaValLeuAla 352
 DB 1009 ATGCTTGAGATGTGTGAGAGAGAAAGACCAAGCCATGACCTGCACGCGGTCTTGGCG 1068
 QY 353 GIUeunProArgProArgArgProProLeuGlnTrpArgTYRSerGIUValISerPheMec 372
 DB 1069 GAGCTTCCCCGCGCCGAGGCGCCCGCTGCAAGTGCAGTACTCGAAGTCTCTTCATG 1128
 QY 373 GIUeunAspLYSPheLeuGIUAAPValArgaenGIYIleTYRProLeuMecAsnPheAla 392
 DB 1129 GAGCTGGAACAATTCCTGGAAGATGTCAAGAAATGAATCTACCACTGATGAACCTTTC 1188
 QY 393 AlaThraArgProLeuGLIYLeuProArgValLeuAlaProProProGIUGIuValGlnYs 412
 DB 1189 GCCACTGACCCCTGGGGCTGCCCGGTGTGTGGCCCCCAACCCCGAGAGAGGTCCAAAG 1248
 QY 413 AlaLYSThrProThrProGIUProPheAspSerGIUThraArgLYeValIleGlnMecGln 432
 DB 1249 GCCAAGAGCCCGAGCGCAAGACCTTTGACTCTGAGACCAAAAGTCAATCCAGATGAG 1308
 QY 433 CYAsnLeuGIUAArgSerGIUAAPLYeAlaArgTrpHISLeuThrLeuLeuValIleu 452
 DB 1309 TGCAACTCTGAGAGAGAGAGAGACCAAGCGCGCTGGCACTTCACTCTCTGTGTGCTG 1368
 QY 453 GIUAAPArgLeuHISArgGIUeunThrTYRAspLeuLeuProThraSPSerAlaGlnAsp 472

Db 1369 GAGAGCCGCTGCACCGGACGCTGACCTGCTCCACGACGACGCGCCAGAC 1428
Qy 473 LeuAlaSerGluLeuValHisTyrGlyPheLeuHisGlu-----AspAspArg 488
Db 1429 CTGCGCTCGGAGCTCGGACACTATGCTTCCTCCACGAGCTGGCGCTGGCGGACGCG 1488
Qy 489 MetLeuLeuAlaPheLeuGluSerThrPheLeuTyrTyrArgGlyThrGlnAla 507
Db 1489 ATGAGCTGGCCGCTTCTCTGAGAGCACCCTCTCAATGACCTGGAGCCAGGCC 1545

RESULT 4
AX056416

LOCUS AX056416 3304 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 60 from Patent WO0073469.
ACCESSION AX056416
VERSION AX056416.1 GI:12229123
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE
1 Plozman, G.D., Martinez, R., Whyte, D. and Sudersanam, S.
AUTHORS
TITLE Protein kinases
JOURNAL Patent: WO 0073469-A 60 07-DEC-2000;
Sugen, Inc. (US)

FEATURES
source 1.3304
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2,3e-196 Length: 3304
Score: 2409.00 Matches: 460
Percent Similarity: 98.72% Conservative: 2
Best Local Similarity: 98.72% Mismatches: 0
Query Match: 90.22% Indels: 6
DB: Gaps: 1

US-10-618-941-99 (1-507) x AX056416 (1-3304)

Qy 40 ArgArgGluGlnValAsnGlnGlyAsnMetProGlyLeuGlnSerThrPheLeuAlaMet 59
Db 1 CAGAGAGAGAGGTAACCAAGGAAACATGCGAGGCTTCAGAGCACTTCCTAGCCATG 60
Qy 60 AspThrGluGluGlyValGluValValTyrAsnGluLeuHisPheGlyAspArgGlyAla 79
Db 61 GACACGAGAGAGGGGTGAGGTGCTGTGGAACGAGCTCCACTTCGAGAGCAGAGAGGCC 120
Qy 80 PheAlaAlaHisGluGluGlyValHisGlnThrValPheGluGlnLeuValLeuValAspHis 99
Db 121 TTGGCGGCGCAGGAGGAGATCCAGACCTGTTGAGAGGCTGTGTGTGAGACAC 180
Qy 100 ProAsnIleValIleLeuHisGlyTyrTyrLeuAspThrSerGlnAlaCysAlaArgVal 119
Db 181 CCGAACAATCGTGAAGTTGCAACAAGTACTGGCTGATACCTCTGAGGCTCGCGAGGCTC 240
Qy 120 IlePheIleThrGluTyrValSerSerGlySerLeuGlnPheLeuValGlyThrLys 139
Db 241 ATCTTCATCAAGAGTACGTGTCTCATCGAGGACCTCAAGCAATTCCTCAAAAGACCAAG 300
Qy 140 LysAsnHisLeuValMetAlaAlaArgAlaTyrLysArgTyrCysThrGlnIleLeuSer 159
Db 301 AAGAAACCAAGGCGCAAGAACGCGCGGCTCGAGAGGCTGTGTGACGACGATCTGTCT 360
Qy 160 AlaLeuSerPheLeuHisAlaCysSerProProIleIleHisGlnLeuLeuThrSerAsp 179
Db 361 GCGCTCAGCTTCTCTGACGCTGACGCCCTCCCAATCATTCACGAGAACCTGACACGCGAC 420

Qy 180 ThrIlePheIleGlnHisAsnGlyLeuIleLysIleGlySerValTyrHisArgIlePhe 199
Db 421 ACCATCTTCATTCAGACACAGCGCTCATCATAGATGGCTCCGTGTGGACCGAATCTTC 480
Qy 200 SerAsnAlaLeuArgProThrAlaLeuProAspAspLeuArgSerProIleArgAla 219
Db 481 TCCAAAT-----GCACCTCCAGATGATCTCCGAGGCCCAATCCGCGCT 522
Qy 220 GluArgGluGlnLeuArgAsnLeuHisPhePheProProGluTyrTyrGluValAlaAsp 239
Db 523 GAGCGAGAGAGAACCTTGGAACCTGCACTTCTCCCGAGAGTATGAGAGGTGGCGCAT 582
Qy 240 GlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGlnMetAlaValLeuGlu 259
Db 583 GGGACCGCTGTGACATCTTCTCTTGGGATGTGTGGCTGAGATGGCTGTACTGGAA 642
Qy 260 IleGlnThrAsnGlyAspThrArgValTyrGluGluAlaIleAlaArgAlaArgHisSer 279
Db 643 ATCCAGACCAATGGGACACCGCGGTCAAGAGAGGCCATTGCTCGCGCCAGCACTCG 702
Qy 280 LeuSerAspProAsnMetArgGluPheIleLeuCysGlyLeuAlaArgAspProAlaArg 239
Db 703 CTGAGTGACCCCAACATGGGGAGTTCACTTGTGCTGCTGCGCCGAGACCTGCGCCG 762
Qy 300 ArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGluValHisSerLeuLys 319
Db 763 CGGCTCTGCGCCACAGCTCTCTCTTCCACCGGCTCTTCCAGAGTGACTGCTGTAAG 822
Qy 320 LeuLeuAlaAlaHisCysPheIleGlnHisGlnTyrLeuMetProGluAsnValValGlu 339
Db 823 CTCTGTGGAGCCCACTGCTTCATCCAGCAAGTACTGATGCTGAGATGTGTGGAG 882
Qy 340 GluLysThrLysAlaMetAspLeuHisAlaValLeuAlaGluLeuProArgProArgArg 359
Db 883 GAGAAACCAAGGCACTGACCTGACGCGGTCTTGGCGGAGCTTCCGCGCCCGCCAGG 942
Qy 360 ProProLeuGlnTyrPheArgTyrSerGluValSerPheMetGluLeuAspLysPheLeuGlu 379
Db 943 CCCCCCTGAGAGGGGTGACTGTGGAAGTCTCTTATGAGGCTGAGCAAAATTCCTGGAG 1002
Qy 380 AspValArgAsnGlyIleTyrProLeuMetAsnPheAlaAlaThrArgProLeuGlyLeu 399
Db 1003 GATGTCAAGATGAGATCTACCACTGATGATGACTTTGCGAGCCACTGACCCCGGGCTG 1062
Qy 400 ProArgValLeuAlaProProProGluGluValGlnLysAlaLysThrProThrProGlu 419
Db 1063 CCGCTGTGTGCTGCCCCACCGCGAGAGGTCCAAAGGCCCAAGCCCGAGCGCCAGAG 1122
Qy 420 ProPheAspSerGluThrArgLysValIleGlnMetGlnCysAsnLeuGluLysSerGlu 439
Db 1123 CCTTTGACTCTGAGACAGAAAGGTCAATCCAGATCAGTGCACCTGAGAGAACGAG 1182
Qy 440 AspLysAlaArgTyrHisLeuThrLeuLeuValLeuGlnAspArgLeuHisArgGln 459
Db 1183 GACAAAGCGCGCTGGCATCTCACTGCTTGTGTGTCTGGAAGACCGCTGCAACCGGAG 1242
Qy 460 LeuThrTyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuValHis 479
Db 1243 CTGACCTTACGACCTGCTCCCAAGAGAGGCCCAAGGACCTCGCTCGAGCTGTGAC 1302
Qy 480 TyrGlyPheLeuHisGluLysAspArgMetLysLeuAlaAlaPheLeuGluSerThrPhe 499
Db 1303 TATGGCTTCTTCCACAGAGACGACCGGATGAACTGCGCTTCTTGGAGAGACACTTC 1362
Qy 500 LeuLysTyrArgGlyThrGlnAla 507
Db 1363 CTCAAGTACCGTGGAGCCAGGCC 1386

RESULT 5
HSM801974
LOCUS HSM801974 3538 bp mRNA linear PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA; cDNA DKFZp434P086 (from clone DKFZp434P086);
partial cds.

ACCESSION AL137662
VERSION AL137662.1 GI:6807782
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3538)
AUTHORS Koehler,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (15-JAN-2000) MIPs, Am Kioferepitz 18a, D-82152 Martinsried, GERMANY
COMMENT
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Charite, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp434P086) is available at the RZPD in Berlin. Please contact the RZPD: Reesourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.
FEATURES
source
1..3538
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434P086"
/cissue_type="testis"
/clone_lib="434 (synonym: htes3). Vector pSport1, host DH10B; sites NotI + SalI"
/dev_stage="adult"
1..3538
/gene="DKFZp434P086"
799..1575
/gene="DKFZp434P086"
/note="similarity to C.elegans H37N21.1/K10D3.5"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA870864.1"
/db_xref="GI:6807783"
/db_xref="GOA:Q9NSY0"
/db_xref="UniProt/TrEMBL:Q9NSY0"
/translation="MCALEMAYLEIQNGDTRVTEBEAIIARARSHISDPMREFFICCLADPARPSAHSILFHRVLFVYHSILKLAHCFIOHQYLMPEMVYBEKKMADLAULAEIAPRRPRLQMRYSBVSFMEDKLELVNRNGIYFLNMFATRRPLGELRVLAAPPEE VQAKPTPEPFDSETRKVIQMCNLSERSDARMMHTLLVLEDRLHRQLTYDLPT DSAQDLASELVHYGFLHEDDRMKLAFLSEFLKRYGTQA"
3471..3476
/gene="DKFZp434P086"
3484
/gene="DKFZp434P086"
ORIGIN
polyA_signal
polyA_site
Alignment Scores:
Pred. No.: 8, 27e-193 Length: 3538
Score: 2368.00 Matches: 460
Percent Similarity: 87.17% Conservative: 2
Best Local Similarity: 86.79% Mismatches: 0
Query Match: 88.69% Indels: 68
DB: 8 Gaps: 2
US-10-618-941-99 (1-507) x HSM801974 (1-3538)
QY 40 ArgatrglGlnValaEnglnGlyaEnMetProGlyLeuGlnSerThirPheLeuAlaMet 59
DB 1 CAGAGGAGAGAGAGGTAAACCAAGGAAACATGCGAGGCTTCAGAGCACTTCCTAGCCATG 60
QY 60 AAPTThGlnGlnGlyValaGlnValaITTPaEnglLLeuHiaPheGlyVaPaRgLyVala 79
DB 61 GACAGGAGAGAGGAGGTAGAGGTGTGTGAACGAGACTCCTCGAGAGACGAGAGGCC 120

QY 80 PheAlaAlaHiaGlnGlyValaIegInThrValaPheGlnGlnLeuValaLeuValaPheHis 99
DB 121 TTCGGCGGCGCAAGAGAGAGATCCAGACCGGTTCGAGCACTGGTGTGGTGAACAC 180
QY 100 ProaEnlLeValaLLeuEnHiaLysTrTTPLeuAAPTThSerGlnAlaCysAlaArgVal 119
DB 181 CGAACATCGTGAAGTGTGACAAAGTACTGGCTGATACCTCTGAGGCGCCGAGGGTTC 240
QY 120 IIPheIleThGlnGlyValaSerSerGlySerLeuLysGlnPheLeuLysTrLys 139
DB 241 ATCTTCATCACAAGATGATCGTTCATCAGGACGCTCAAGCAATCTCTCAAAAGACCAAG 300
QY 140 LysaEnHiaLysValaMetAsnAlaArg----- 148
DB 301 AAGAACCAAGGCAATGAAGCCGGGATATGGGAGCGGCGCTGGGAGCAGCAGGGGAC 360
QY 148 ----- 148
DB 361 AGAGCGGGTGGGCGAGCTCGGGAGTGGAGTGAAGGGGTGCCCGGCGCTCGG 420
QY 148 ----- 148
DB 421 ACAGGGGCTGGCGAGATCGGGGCGGGGCTCGAGGCCCGGCTCTCGGCGCC 480
QY 149 -----AlaTrpLysArgTrpCysThrGlnIle 157
DB 481 ACCGACCGAGCGAGATCGTGCCTCGCGCGCCAGGCTGGAAGCGCTGGTGCAGGAGATC 540
QY 158 LeuSerAlaLeuSerPheLeuHiaLacySerProProIleIleHiaGlyAsnLeuThr 177
DB 541 CTGTCTGGCGCTCAGCTTCCTGCAAGCCCTGCAAGCCCAATCATCAGCGGAACCTGACC 600
QY 178 SerAAPTThIlePheIleGlnHiaPheGlyLeuLysIleGlySerValITPhHisArg 197
DB 601 AGGACACCATTTATTCAGCAACAGCGCTCATCAAGATGGCTCGCTGGACCGCA 660
QY 198 IIPheSerAsnAlaLeuArgProProThraLeuProAspAspLeuArgSerProIle 217
DB 661 ATCTTCTCAAT-----GACTTCCAGATGATCTCGAAGCCCAAC 702
QY 218 ArgAlaGlnValGlnGlnLeuLysValaLeuHiaPhePheProProGlnTrGlyGlnVal 237
DB 703 CGCGCTGAGCGAGAGAACTTCGGAACCTGCACTTCTCCCCCAAGATGAGAGGTC 762
QY 228 AlaAPGlyThraValaValaAspIlePheSerPheGlyMetCysValaLeuGlnMetAlaVal 257
DB 763 GCGATGGGACCGCTGTGACATCTTCTCTTGGAGATGTGCGCTGGAGATGCTGTGA 822
QY 258 LeuGlnIleGlnThraEnglyVaPThraArgValThrGlnGlnAlaIleAlaArgAlaArg 277
DB 823 CTGGAAATCCAGACCAATGGGGACACCGGGTCAAGAGAGGCCATTCCTGGCCGACG 882
QY 278 HisSerLeuSerAAPTProAspMetArgGlnPheIleLeuCysCysLeuAlaArgAAPTPro 297
DB 883 CACTGCTGAGTGAACCCCAACATGCGGGAGTTCACTTGTGCTGCTGCGCCGAGACCT 942
QY 298 AlaArgArgProSerAlaHisSerLeuLeuPheHiaArgValaLeuPheGlnValaHisSer 317
DB 943 GCCCGCGCGCTCTGCGCACAGCTCTCTTCCACCGCGTCTCTGAGGTGACCTTG 1002
QY 318 LeuLysLeuLeuAlaAlaHiaCysPheIleGlnHiaGlnTrLysLeuMetProGlnLysVal 337
DB 1003 CTGAAGCTCTCTGGAGCCCACTGCTTCATCCAGCAACGATACCTCATGCTTAGAAATGTG 1062
QY 338 ValGlnGlnLysTrLysValaMetAspLeuHiaLysValaLLeuAlaGlnLeuProArgPro 357
DB 1063 GTGAGAGAGAGAGACCAAGGCATGAGCTGCAAGCGGCTTGGCGAGCTTCCCGGCCC 1122
QY 358 ArgArgProProLeuGlnITPhArgTrSerGlyValaSerPheMetGlnLeuAspLysPhe 377
DB 1123 CGCAGGCCCGCGTGCATGGCGGTAATCGGAAAGTCTCTTCATGAGCTGGACCAAAATTC 1182
QY 378 LeuGlnAAPTValaArgAsnGlyIleTyProLeuMetAspPheAlaIleThraArgProLeu 397

Dd		1183	CTGGAGGATGTCCAGGAATTGCAGATTCACCACTGACTGAACTTTCAGCACCCTTG	1242
Oy		398	GlyLeuProArgValIleuAlaProRobroPugluValIglnLysAlaIstheProThr	417
Dd		1243	GGCGTCCCCCGTTGCTGCCGCCCACTCCGAGGAGGTCCAATAAGCCAAAGCCCCGAGC	1302
Oy		418	ProgluProPhaeAspSerClunThrxTrglyrsValIIlegImetGlncYasAmleugiUarq	437
Dd		1303	CCAGAGCCCTTTGACTCTTGAGACCAAGAAGTCAATCCAGATGCACTGCACAACCTGGAGGA	1362
Oy		438	SerGIuaAPLyALeArGTphIsleuthrLeuleuLeuValIleuGIuaSPArghLeuhIs	457
Dd		1353	ACCGAGCACAGGCCCGCTGGCATCTCACTCTGCTTCTGGTGTGGAAGAACCGGCTGGAC	1422
Oy		458	ArgGInLeuthrTYrAspLeuleuBProThrxApSeraIagiNaSpLeualAserCIudeu	477
Dd		1423	CCGCGAGCTGACCTTAGACCTGCTGCCAACGACAGCGCCAGAGACTGGCTCCGAGCTC	1482
Oy		478	ValHsrTyrgLYpheLeuHiBgIUaSPArghMetIlysleuAlalaPheLeugiUser	497
Dd		1483	GTGCACTATAGCTTCTCCACGAGGACGACCGGATGAAGCTGGCGCTCTTGAGAGC	1542
Oy		498	ThrPheLeuLyTYrTYrArgIGlYThrcgInla	507
Dd		1543	ACCTTCCTCAAGTACCGTGGAGCCAGGCC	1572
RESULT 6				
ARJ10264				
LOCUS	ARJ10264	981 bp	DNA	linear PAT 12-JUN-2003
DEFINITION	Sequence 9 from patent US 6558903.			
ACCESSION	ARJ10264			
VERSION	ARJ10264.1			
KEYWORDS	GI :31703044			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	I (bases 1 to 981)			
TITLE	Hodge,M.R.			
JOURNAL	Kinases and uses thereof Patent: US 6558903-A 9 06-MAY-2003; Millennium Pharmaceuticals, Inc.; Cambridge, MA			
FEATURES	Location/Qualifiers			
Source	1..981 /organism="unknown" /mol_type="genomic DNA"			
ORIGIN				
Alignment Scores:				
Pred. No.:	4,12e-129	Length:	981	
Score:	1617.00	Matches:	312	
Percent Similarity:	96.89%	Conservative:	0	
Best Local Similarity:	96.89%	Mismatches:	3	
Query Match:	60.56%	Indels:	7	
DB:	6	Gaps:	1	
US-10-618-941-99 (1-507) x ARJ10264 (1-981)				
Oy		1	MetaIAAIaPProgiUpProAlaPProArgArghIaaArgGUaRGUaRGUaRGUaRGUaSP	20
Dd		35	ATGGGGGGCCCAGAGCCCGCGCCCAAGCGGGGCCGGGAAACGGAGACGGGAGGAC	94
Oy		21	GlueSerGIuaBPgluSerASPILeleugiUGluSerProCySGIYAgtTPGLnLyARG	40
Dd		95	GAGAGCCAGACGAGAGGACATCTCTGAGGAAAACCCGTGTGCTGCTGGCAAACGA	154
Oy		41	ArgGUgInVaIaengIngIYaSmweBProGIyeuInserThrpHeulaIameASp	60
Dd		155	C GGAGGACGCTAAACCAAGGAAACATGCAAGGCTTGAGACCACTTCATGACATGAC	214
Oy		61	ThrgIuGiUGluGYalGIuValValITrrPaengIuLeuHisPrnegIyaSPA-glyrsAlaPhe	80
Dd		215	ACGGAGGAGGGGGGTGAAGTGTGTGGAAACAGCTCCACTTCGGAGACAGGAAGGCTTC	274

QY	81	AlaAlaIhIeGIuGInLyIleGIInThrValPheGIuGInLeuValLeuValAspIlePro	100
Db	275	GGGGGCGACGAGAGAGAATCCAGACCGTGTTCGAGACGCTGTGCTGTGGACCAACCG	334
QY	101	AsnIleValIyLeuIhIeLyTyTTrpleuAspThrSerGIuValCyAlaArgValIle	120
Db	335	AACATCGTGAAGTTCACAAAGTACTGGCTGGATCTCTGAGGCCCTGGCGGAGGTCAATC	394
QY	121	PheIleThrGIuTYrValSerSerGIySerIleuLyGIInPheLeuIyLyThrIyAlLyS	140
Db	395	TTCAATCAACAGAGTACGTGTCAATCAGGACCGCTCAAGCAATTCCTCAAAAAGCCAAAGAG	454
QY	141	AsnHISlySAIleMeCAsnAlaArgAlaTrpLySAgTAPCySthGIInIleuSerAla	160
Db	455	AACCAAGGCGCATGAAAGCCCGGGCTCGAAGCGCTGTGTGACCAAGATCCTGTCTGG	514
QY	161	LeuSerPheLeuHISAlaCySerProGIleIleHISGIyAsnLeuThrSerAspThr	180
Db	515	CTCAGCTTCCTGACAGCGCTGACGCCCCCATCATCCAGGGAACTTGACCAAGCAAC	574
QY	181	IlePheIleGIInHISArgGIyLeuIleuYsIleGIySerValTrpHISArgIlePheSer	200
Db	575	ATCTTCATTCAGCAACAGCGCTCATCAAGATCGGCTCCGTGTGGCAACCGAATCTTCTCC	634
QY	201	AsnAlaIeUArgProTrpThrAlaIeUProAspAlaIeUArgSerProIleArgAlaGIu	220
Db	635	AAT-----GCACTTCAGATGATCTCCGAAAGCCCATCCGCGCTAG	676
QY	221	ArgGIuGIuLeuArgAsnLeuHISPhePheProGIuTYrGIyGIuValAlaAspGIy	240
Db	677	CGAAGGAACCTTCGGAACCTGCACTTCTCCCCCAAGATGAGAGAGTGGCGGATGG	736
QY	241	ThrAlaValAspIlePheSerPheGIyMeCyAlaIeUGIuMeAlaValIeUGIuIle	260
Db	737	ACCGGTGTGACATCTTCTTCTTGGGATGTGTGGCTGGAGATGGCTGTACTGGAATTC	796
QY	261	GIInThrAsnGIyAspThrArgValThrGIuGIuAlaIleAlaArgAlaArgHISerIleu	280
Db	797	CNAGCAATGGGGGACACCGCGGTCAACAGAGGCCATTGCTCGGCGCAGGCACTTCGGT	856
QY	281	SerAspProAsnMeCArgIuPheIleuCyAlaIeUAlaArgAspProAlaArgArg	300
Db	857	AGTGACCCCAACATGGCGGAGTTATCTTGTGCTGGCGCCGGGACCTGGCCGAGG	916
QY	301	ProSerAlaHISerIleuLeuPheHISArgValIeUPhelGIuValHISerIleuLyLe	320
Db	917	CCCTCTGTCCACAGCTCTCTCTTCACACGCGTGTCTTNGAGGTGCACTCGCTGAAGCT	976
QY	320	uLeu 321	
Db	977	CCTG 980	
RESULT 7			
LOCUS	AR640307	981 bp	DNA
DEFINITION	Sequence 9 from patent US 6858418.		
ACCESSION	AR640307		
VERSION	AR640307.1	GI:62774642	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 981)		
AUTHORS	Hodge,M.R.		
TITLE	Kinases and uses thereof		
JOURNAL	Patent: US 6858418-A 9 22-FEB-2005;		
FEATURES	Millennium Pharmaceutical, Inc.; Cambridge, MA		
SOURCE	Location/Qualifiers		
	1..981		
	/organism="unknown"		
ORIGIN	/mol_type="genomic"		

AUTHORS

Birken, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Bouglavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chepel, Y., Collymore, A., Cook, A., Cooke, P., Cornu, B., Dattellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, J., Grand-Pierre, N., Hafez, N., Haggopian, D., Hages, B., Hall, J., Horton, L., Holme, W., Iliev, I., Johnson, R., Jones, C., Kamet, A., Karakas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramamany, U., Raymond, C., Retta, R., Riese, C., Rogov, P., Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Teefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL

COMMENT

Submitted (02-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 2, 2003 this sequence version replaced gi:28827929.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: L23072

Center clone name: 299_M_14

FEATURES

source

Only the middle 92.3 kilobases of this clone are being submitted.
The remainder overlaps either accession number AC109322 [WICGR project L23183]
or accession number AC105219 [WICGR project L23081].

Location/Qualifiers

1. 92291

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="8"

/map="8"

/clone="RP11-299M14"

/clone_lib="RP11-299M14"

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

/rpt_family="L1MD2"

7326..7349

/rpt_family="(CAAAA)n"

complement(7377..7458)

/rpt_family="L1MD2"

8091..8207

/rpt_family="MIR"

8208..8505

/rpt_family="AluSg"

8506..8580

/rpt_family="MIR"

complement(8785..9127)

/rpt_family="AluJo"

9225..9304

/rpt_family="(ATGCTG)n"

9420..9660

/rpt_family="AluSg"

9661..10614

/rpt_family="L1R5_Hs"

10615..10690

/rpt_family="AluSg"

10697..10994

/rpt_family="AluY"

11474..12045

/rpt_family="L1MED4"

complement(12065..12372)

/rpt_family="AluY"

12379..12519

/rpt_family="L1MED4"

complement(12918..13211)

/rpt_family="AluSx"

complement(13213..13346)

/rpt_family="AluSg/x"

complement(13363..13736)

/rpt_family="MER87"

complement(14215..14374)

/rpt_family="MLTIF2"

14698..15000

/rpt_family="AluJo"

complement(15346..15495)

/rpt_family="MLTIC"

complement(15500..16464)

/rpt_family="L1R13"

complement(16472..16738)

/rpt_family="MLTIC"

16783..16874

/rpt_family="L2"

complement(16869..17347)

/rpt_family="L2"

complement(17349..17752)

/rpt_family="MER77"

17753..17956

/rpt_family="MER30"

complement(17957..18089)

/rpt_family="MER77"

complement(18519..18684)

/rpt_family="L2"

19020..19200

/rpt_family="G-rich"

19204..19235

/rpt_family="GC-rich"

19295..19344

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

3.9e-126

1612.00

28.84%

28.79%

60.37%

8

US-10-618-941-99 (1-507) x AC105049 (1-922291)

92291

498

1

8

1231

17

Db 21646 AAGCAGTTCCTGCTCCCTCCAGCTGCTACCGTGTCAAGGTAGAGGCTAGCAAGGT 21705
 QY 286 ----- 286
 Db 21706 GCTCCAGGACAGCTGGAGAACATCACAGTTTGTGAGTGGGGTATACAGCTGAC 21765
 QY 286 ----- 286
 Db 21766 TGACCCCTCCCAAGAGAGACCTGAGTGGAGCATACTAGTTTCCAGGCTTGAAGGCC 21825
 QY 286 ----- 286
 Db 21826 TAAGGGTTCAGACATTTCTAAGACAAGAGGCAAGAGACCTCCCTGTGTGCC 21885
 QY 286 ----- 286
 Db 21886 ATACCATGCCCGTTTATTAAGAAACCAAGTTTCAAGAGATGAGTCTTTGCCAAG 21945
 QY 286 ----- 286
 Db 21946 GCTTTAGATGATGATGAGCCAAATCTGCACTGGGTCTGCTTCACTTCAAGTCTTT 22005
 QY 286 ----- 286
 Db 22006 GACTTTAAGCTTCAGAGAGTATGAGGCTTCAGCTTGTGGGGAGATGTTGGTCCA 22065
 QY 286 ----- 286
 Db 22066 AGCAGGAAAAAGATAGGGATGAGAGCTGCATCCAGCCCCCTGTGTGCGAGTCT 22125
 QY 286 ----- 286
 Db 22126 AGACTTAATCTCATGAGCTCAGAGATGGGGGTCTGTGTGGGGCTGTGGCGAGCTGG 22185
 QY 286 ----- 286
 Db 22186 AGCGGGCTGGGACGAGATGAGACAGATGGGGGACCCATCAAGTGTCTCCAGTGTGG 22245
 QY 286 ----- 286
 Db 22246 GGGCCGAGTGAAGCCGATCAAAAGGCCATGGGTAGAGGCTTGCAGAGTGTGATTTAGGA 22305
 QY 286 ----- 286
 Db 22306 TCATGAGCTTCTCCCAAGATTTTGTGGGGTCCAGGAGAGACATGTTCTCCATTC 22365
 QY 286 ----- 286
 Db 22366 CCAAGAACCAAGCTGTCTCAAGGTCAGAGGACCAAGTCTGGCTGGCTGTCTGTGG 22425
 QY 286 ----- 286
 Db 22426 GACTCTCTCTATGGGATCCTGATGCTGTGTGTCTGCTGAGAGAGACATGGGGAA 22485
 QY 286 ----- 286
 Db 22486 CTGAGGGGCACTTCTCTGAACTGGGGGCCAAGACAGAGACCTTGCACACCA 22545
 QY 287 ----- 287
 Db 22546 GCGTCTCTCCCTGTGGCCCAAGAGTTTCTTTGTGTGTGGCCCGGACCTTGGCCG 22605
 QY 299 gATgPrOsErAlhIsSerLeuLeuPhehIaRgValLeuPheGluValhIsSerLeuLy 319
 Db 22606 CCGGCGCTTGGCCACAGCTCTCTTCCACCGCGTGTCTTGAAGGTGACATCGTGAA 22665
 QY 319 sIleuLeuAlaAlaHIsCyArPheIleGlnHIs----- 329
 Db 22666 GCTCTGTGGAGCCCATGCTTCAATCCAGCACAGTGTGACAGACAGCGCGCGGGGG 22725
 QY 329 ----- 329
 Db 22726 TGGCTGGGAGGATGAGGTTGGGACGTGAAGGGGCCCTGCAAGAACCGGCTGAGCA 22785

QY 330 -----GlnTyrLeuMetProGluAsnValValGluGluLy 341
 Db 22786 CACGCCCTGATGCTCTTGTGGCCGACAGAC-CTCATCTCTGAGATGTGTGGAGAGAA 22844
 QY 341 sThrLyAlaMetAspLeuHIsAlaValLeuAlaGluLeuProArgProArgProPr 361
 Db 22845 GACCAAGGCCATGACCTTGACCGCGATCTTGGCGGAGCTTCCCGGCCGACAGGCCCC 22904
 QY 361 OleuGlnTyrArg----- 365
 Db 22905 GCTGAGTGGCG-GTGAGAGTCAGACCCAGAAAGCTCCCTCCACACATCCCTCC 22963
 QY 365 ----- 365
 Db 22964 CCGCAGCTCCCTCCCTCCGCCCCGACAGCCCTCAATGCTCCCATTCATGCCAG 23023
 QY 366 -TyrSerGluValSerPheMetGluLeuAspLysPheLeuGluAspVal----- 381
 Db 23024 GTACTCGAAGTCTCTTCAATGAGCTGACAAATTCCTGAGAGATGT-CAGGTGAGAAC 23082
 QY 381 ----- 381
 Db 23083 TGATGTAAAGTTGGGGGCAAGCAGCTGTCAAGGCTGAACCCCAAGCCACTGTGTGT 23142
 QY 382 ----ArgAsnGlyIleTyrProLeuMetAsnPheAlaIleThrArgProLeuGlyLeuPr 400
 Db 23143 CCCAGGAATGGAATCAACCACTGATGAATTTGGCAGCCACTGACCCCTGGGGGTGCC 23202
 QY 400 oArGValLeuAlaProProProGluGluValGlnLysAlaIleThrProThrProGluPr 420
 Db 23203 CCGTGTGCTGGCCCAACCCCGAGAGAGTCCAAAAGCCAAAGCCCGACGAGAGCC 23262
 QY 420 oPheAspSerGluThrArgLysVal----- 428
 Db 23263 CTTGACTCTGAGACCAAGAAAGGT-GAGTCCCTCTCTGTCTGATGCCATCCCTCA 23321
 QY 428 ----- 428
 Db 23322 GACTTTCAACAGCAGATTAAGAGTGCCCTCTCTTCTTAAAGGCTGATCCCAAG 23381
 QY 428 ----- 428
 Db 23382 CCAAGCTCTCTCTGTGCTGCCAAGAACAGATCCAGGCTGACCCCTGTGTGGGGTCA 23441
 QY 428 ----- 428
 Db 23442 GGGGGTTATAGACAGAGCTGGGCACTCTCTTTCCGGAGAGCTCCCTGTCCCTCAGAT 23501
 QY 428 ----- 428
 Db 23502 CCTTCCCTGATCTGTGCTGTGCTGCTGAACCTACTCCAGAGATCTTCTCCAGAACT 23561
 QY 428 ----- 428
 Db 23562 CACCCGGGCTTTCACACCTGCGACGAATCTTCCCAACCCCGCACCCCTCCACCCCC 23621
 QY 428 ----- 428
 Db 23622 CCCCCACCCCTCCATCCCCCACTGTGTGATTTCCCGGACATATATGCTGTGTGTCT 23681
 QY 428 ----- 428
 Db 23682 CCCAGACTTTTCTGACAGACACCCCTCTGCGCAGGGGCTCTCTCGGTGAGTGAGC 23741
 QY 428 ----- 428
 Db 23742 CAGAAATTCATGTCTTAACAGGCAAACTGCTGTCTGTGCTGCGACAGAGATCATGGGA 23801
 QY 428 ----- 428
 Db 23802 GTCTCAGCTGACGCCAGGCTCTGGGAGGCTGATTCGGCAGAGCACCGTGCATTGGC 23861

OY	428	-----	428
Db	23862	GCGGCTCTGTCGCGCCAACTCCCTTTCCCAAGCCCAAGTTCCCTTTGTTCAGT	23921
OY	428	-----	428
Db	23922	TGTTCTCTCGCGGGTGAAGCTCTATTTCAGCTTACGTACCCAGCCCTAATGTGACT	23981
OY	428	-----	428
Db	23982	CTTCTCGAAGCTTCTTAGCTTGCCAAAGCCTAGCAGTAGATGGCTCTTAGAG	24041
OY	428	-----	428
Db	24042	ATTCCTCCCTTTCCAGTCTTTGCTGTCCCTTCCGCGCCCTTTGGGGCGCGGCAATGG	24101
OY	429	-----	433
Db	24102	TCTTGCTGCTGGGTTTGAGTCCCAAGCTTCCCGCTGACGTATCATTCAGATGCAGTG	24161
OY	433	saaluEgluarGserGluAplysalargTyrHis-----	445
Db	24162	CAACTGGAGAGAAAGCGAGACAGCGCGCTGCATGTAGCGGGGCTGGAGAGCGG	24221
OY	445	-----	445
Db	24222	CGTGCCTGGGGAAGCGGAGCGGTGCGGAGCCCTGAACCCCAAGCCGACAGCCCC	24281
OY	446	-----LeuThrLeuLeuLeuValIleuGluAspArgLeuHisArgGlnLeuTh	461
Db	24282	GCCGCTCTCCAGCTCACTGTCTTTGTGTCTGTGAAGAACCGGCTGCACCGGAGCTGAC	24341
OY	461	TTYAspLeuLeuPro-----	466
Db	24342	CTAGACCTGCTCCCAAGTAGAGCTGGGAGGGGACTTGGGGGGCGGCGGTGACCCOAGG	24401
OY	467	-----ThrAspSerIleG	471
Db	24402	CGCGCGGTGGCAGCGGCTCATGCTCGCTTACGCAATTCCTTCAGCGGACAGCGCCC	24461
OY	471	IAspLeuAlaSerGlnLeuValHisTyrGlyPheLeuHisGlu-----	485
Db	24462	AGGACTCTGCTCGGAGCTCGTGACCTATGGCTTCTTCCACGA-GGTGCGCTGGCGGCTG	24521
OY	485	-----	485
Db	24521	CGGCGCGGCTGGGCGAGGGGGCCACGCGGCGAGCGCGCCCTCTCCCTCCCATGGCT	24581
OY	486	-----AspAspArgMetIleValAlaIleLeuGlnSerThrPheLeuI	501
Db	24581	CCCTCTCTCCGACAGGACGACCGGATGAAGCTGGCGGCTTCTGGAGAGCACTTCTCTCA	24641
OY	501	yeTYArgGlyThrGlnAla	507
Db	24641	AGTACCGTGGGACCCAGGCG	24660
RESULT 9			
AC116487	174949 bp	DNA	linear
LOCUS	Mus musculus chromosome 15 clone RP23-42819 map 15, ***	HTG 02-AUG-2005	
DEFINITION	Mus musculus chromosome 15 clone RP23-42819 map 15, ***	SEQUENCING	
IN PROGRESS ***.			
AC116487			
AC116487.13	GI:71648610		
VERSION	HTG: HTGS PHASE2; HTGS FULLTOP; HTGS_ACTIVEFIN.		
KEYWORDS	Mus musculus (house mouse)		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 174949)		
AUTHORS	Birren, B., Nusbaum, C. and Lander, E.		
TITLE	Mus musculus chromosome 15, clone RP23-42819		
JOURNAL	Unpublished		

REFERENCE
AUTHORS

2 (bases 1 to 174949)
Biren, B., Linton, L., Nusbäum, C., Lander, E., All, A., Allen, N., Anderson, S., Barna, N., Baehten, V., Bloom, T., Bogunlavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R., Landers, T., Leoczky, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Mayor, J., Marcini, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Minova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Ntshu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunthang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schochack, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Teefaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (28-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 174949)
Biren, B., Nusbäum, C., Lander, E., Abouellell, A., Allen, N., Anderson, M., Anderson, S., Archchi, H.M., Barna, N., Baehten, V., Bloom, T., Bogunlavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagoian, D., Jones, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., MacDonald, P., Mayor, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Minova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunthang, P., Pierre, N., Rachupka, A., Ramaany, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schochack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Teefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (02-AUG-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 2, 2005 this sequence version replaced gi:17067273.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Broad Institute of MIT and Harvard
Center code: WTB8
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: L23947
Center clone name: 428_1_9

NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES

Bouyce

```

1 174949: contig of 174949 bp in length.
Location/Qualifiers
1. 174949
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="15"
/map="15"
/clone="RP23-42818"
/clone_lib="RPC1-23 Female Mouse BAC"

```

ORIGIN

Alignment Scores:

Pred. No.:	1.29e-121	Length:	174949
Score:	1553.50	Weight:	100

Score:	1563.50	Matches:
Percent Similarity:	32.50%	Conservative:

Best Local Similarity: 31.69%
Mismatches:

Query Match:	58.56%	Indels:
--------------	--------	---------

DB: 14 Gaps:

US-10-618-941-99 (1-507) X AC116487 (1-174949)

US-10-618-941-99 (1-507) x AC116487 (1-174949)

Oy		1	MetAlaAlaProGluProAlaProArgArgLaaArgLuuArgLuuArgGluuAsp	20
Dd		5486	ATGGCGGCCCGAGACCAGCCGCAGAGAAGC-----CGAGACCGCAACCGGAAGAC	5539
Oy		21	GluSerGIuaerGIuSerAepIleuengluiuserProCySGIyaGrPgluIyvarg	40
Dd		5540	GAGAGCGAGCACGAGAGCATCTGTGAGAGAGCCCGTGTGCCGTGGCAGAACGG	5599
Oy		41	ArgGluiNvaI-----	44
Dd		5600	CGAGAGCAGGT-ggcACGAGTGCGAGTGAACGCCGCTCCGTACTAGATAGAAAG	5658
Oy		44	-----	44
Dd		5659	CCATTGGATCCGTGGGCGAGCCAATAAGAGGTTCGCCGAGACCCCTGGACTGA	5718
Oy		44	-----	44
Dd		5719	GGCCCTGGATCCCTTTGGCCCCCAAAGCCCTGTGAGAAACCCCATTCCTTGAACA	5778
Oy		44	-----	44
Dd		5779	TGTCTACTGGGACCACTCCACGNGCCTTTGATCCCTACTAACCGCTCGGGATTC	5838
Oy		44	-----	44
Dd		5839	CCCAACCTTAAMACMGAGTACCTTGAGCCAGCCCCTGAGACACACCTTCTCCCC	5898
Oy		44	-----	44
Dd		5899	TGGATTGTGATTTCCCAAGCTTAACTCTCAAGCCCTCCCACTGAGTGGGCAAGCTAAC	5958
Oy		44	-----	44
Dd		5959	AGCATTTCAACGCTGGGGGAAGTTGGCGGAGACCAAGTCCACACCCCTGGTCTATT	6018
Oy		45	-----AnglgiIyAnMetProGIyLeuGIu	53
Dd		6019	CGATACGACACACATTTCTCCACCTAGGTGAACCAAGGAACTAGCCCGGATCCAG	6078
Oy		54	SeThrPheluIameIeArPhrThrgluGIuGIuValIuTranGIuIeuHIs	73
Dd		6079	AGCACTTTCTGGCCATGAGACAGAGAGAGGTGTAAAGTGTGTGAACGAACCTTCA	6138
Oy		74	PheGIyAephrglyvalaphelaIaIHISGIu-----	84
Dd		6139	TTGGGGATAGAAAGCCTTTGGGGCCCATGA-gtGAGGCTGTCTATACCCGCTTGC	6197
Oy		85	-----GluYvIIegInThrVal	90
Dd		6198	CCCATACCTTGATTGGCCCCCAACACATCTCTTCCCGAGAGAGATTCAGACCATG	6257

QY	91	PheGIuGI nLeuValLeuValAspHisProAsnIleValIleuHisIleValTyrTrrLeu	110
Db	6258	TTTGTGACAGCTGGCCGCTGTGTGACCATCCCAACATGTCAGCTGCACAAATGTTGGCTG	6312
QY	111	AspTrsSerGIuAla	115
Db	6318	GACGCGCTCTGAAGC-CGCGCGAGGGTGAAGCTAGGGGGGGGTGAAGGGCGAGTGAAGC	6378
QY	116	-----CysAlaArg	118
Db	6377	ACAGCGTGGTGGGCGACATGAGACCCACCGGACAGCGGAGCCCGCTGGCCCGGAG	6438
QY	119	ValIlePheIleThrGIuTyrValIserSerGIysIleuIysGIuPheLeuIlyIysIleThr	138
Db	6437	GTCATCTTCAACACAAATACGATGTGCTGTGGCAGCTTCAGACAGTTCCTCAAAAAGACC	6496
QY	139	LysIlyAsnHisIlyValIleAsnAlaArg	148
Db	6497	AAGAAGAACCAAGGCGCATGAGCGCCGGGTATGGGAGGTGGGTGGGACTGGAAGGA	6555
QY	148	-----	148
Db	6557	GTCACAGCAACTTAGGAGACAGTAGTGGGGCGGGGGACTCCGAAAACTTCGATTACGA	6618
QY	148	-----	148
Db	6617	GTGGGGTAGGGGGGGGGCTCCGACAGCTCAGACAGCTTTTGTGTGCCACCCGAGGAGTC	6678
QY	149	-----AlaTrrPlySerGIyTrrCysThrGIuIleLeuSerAlaLeu	161
Db	6677	GTGCGTCCGCTGCAGGCGCTGGAAGCGCTGTGATCGAGATCCGTCTGCATCAGGGT	6738
QY	161	-----	161
Db	6737	AGAGCTTGCGCTTGCCACAGCGCTTGCCCTGCTGCACCTTTCCGCTGCTTGAC	6796
QY	162	-----SerPheLeuHisIAlaCysSerProPr	170
Db	6797	CCGGTCTTCTCGCTCATCTCATCTCTAGTGTGGAGTTTGTGACGCGCTGCATCCGCC	6856
QY	170	GIleIleHisGlyValnLeuThrSerAspThrIlePheIleGIuHisIAsnGlyLeuIleIy	190
Db	6857	CATCATTCACCGGAACTGACAGACGACACCATCTTCATTCAGACATAGCGCTCATCAA	6916
QY	190	AlleGIySer	193
Db	6917	GATCGGCTCCGGTCTGGCCAGAGCGGGGACAGGGGGGACAGGGGTGGGAACTTGC	6976
QY	193	-----	193
Db	6977	GGAGACCTGACAGAGACTGCCAGTCTCCCGCCCACTGACATCTGTCTTCTCTC	7036
QY	194	-----ValTrrPheIAsnGlyIlePheSerAsn	201
Db	7037	TGCGTTGCGTGTGCAGGTGTGTACCGCATCTTCTCAAAAGTGTGTGGGTCTCTGGGG	7096
QY	202	-----AlaLeuArgProPro-----ThrAlaLeuProA	211
Db	7097	GTGTGGGGAGCCATGGAGAACTGTCTGTTAGTCCACCTTTCTCTCAGCACTTCTG	7156
QY	211	SPAspLeuArgSerProIleArgAlaGIuArgGIuLeuArgAsnLeuHisPhePheP	231
Db	7157	ATGATCTCCCGAGCCCTATTCGACTGAACGGGAAAGAACTCCGAAACCTGACATTTTCC	7216
QY	231	roProGIuTyrGIy	235
Db	7217	CACCAAGATATGGCGGTGATGCTCCTAGCTGCTCTCAGGATCTGACGGGCTAGCACATC	7276
QY	236	-----GluValAlaAspGIyTrrAlaVal	243
Db	7277	AGAACTGACCAAGCTGTCTTGTCTTCCCATTTCCAAAGTCAATGATGTGGAGCTGCTGTG	7336

QY 244 Asp1ePheSerPheGlyMetCysAlaLeuGlu----- 254
Db 7337 GAAATCTTCTCTTGGGATGTGTGCACTGGA--GTAATCTCACTCTCTCAATAGCCC 7395
QY 255 -----MetAlaVal 257
Db 7396 TGCTCCAGGGCTACGCCATCCACCAATTCCAATTGTCTCTGACCACCGAGTGGCTGTA 7455
QY 258 LeuGluIleGlnThrLeuGlyAspThrArgValThiGlnGluAlaIleAlaArgAlaArg 277
Db 7456 CTGAGATCTCAACCCAAACGGGATACCAAGTACAGAAAGAGGCCATCGCTGAGCCAGG 7515
QY 278 HisSerLeuSerAspProAspMetArg----- 286
Db 7516 CACTCACTGATGATCCCAACATGCGGGTACGACGACTTACCTTGCCCTTGGGGCTGGC 7575
QY 286 ----- 286
Db 7576 TGCAGGTGCTTCAAGCTTACCTTGAAGTGCTTAGGACAGCCAGAGGGGAGCCTGGTGC 7635
QY 286 ----- 286
Db 7636 GTAATCTGAGAGGGGAGCCTGTGTACATCTGAGAGTACAGAGAGCTAAGGC 7695
QY 286 ----- 286
Db 7696 CCACCTGAAAGAAAGACTCTTCCAGAAAGCTCACAGACTCCCAAGGGTAAATGAAAGTAG 7755
QY 286 ----- 286
Db 7756 GCATAGCTGTGTTGAGAGCCAAAGGCTTGCTGTCTCAAGTGTCCAGAAATCAGATCTTAG 7815
QY 286 ----- 286
Db 7816 CTGCTAATGGATGACTTCTGTCTCATATCAGAAAGGCTTTTCAGCATCCTTAG 7875
QY 286 ----- 286
Db 7876 GTGGAGATGTGTGATTTGGACCGGGGTCTAAGATTTGAAGACCTCTCTAGAG 7935
QY 287 -----GluPheIleLeuCy 291
Db 7936 CGAGACTTTTTCAGCTATCACTGTCTTTCTTGTTTCAACCTAAGAAATTCATCTCTTC 7995
QY 291 sCySLeuAlaArgProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgVal 311
Db 7996 CTGCTGTGGCCCGGACCTGTGCGCGACCTTCAGGCCACAACTCTCTTTCACCGAAGT 8055
QY 311 IleuPheGluValHisSerLeuIysLeuLeuAlaAlaHisCyPheIleGlnHisGln-T 331
Db 8056 GCTCTTTGAGGTGCACTCGTGAAGCTGTGGACGCTCACTGCTTCAATCCAGCACAGTGG 8115
QY 331 yr----- 331
Db 8116 TGAGGGGAGAAAGGCGGGGTGGGTACAGGATCTGGGGTCAAGTATGGAAAGGTCTTT 8175
QY 332 -----LeuMetP 334
Db 8176 GCAGAGCCCTTAGATTGGGTCCATCCCATAGACACTTGTGTCTTACACAGACTCATGCT 8235
QY 334 roGluAsnValValGluGluIysThrIysAlaMetAspLeuHisAlaValLeuAlaGlu 354
Db 8236 CTGAGATGTGTAGAGAAAGAACCAAGGCCATGACCTTCATGCAAGTTTGGCTGAGG 8295
QY 354 euProArgProArgArgProProLeuGlnIlyArg----- 365
Db 8296 TGCACGAGCCCATGAGACCCCAATGACAGTGGCG--GTAGTATGGCCAAACGAGACCTTC 8354
QY 365 ----- 365
Db 8355 CCTTCTAGCTGTGGCTAGGCTTGGCTTGCCTGACATCTCTCATGACTTTCCACACCATTT 8414
QY 366 -----TyrSerGluValSerPheMetGluLeuAspIysPheLeuGluAspVal----- 381

Db 8415 CTAGGACTCAGAGGCTCTCTTCTTGGAGCTGGACAAATTCTTAGAGATGT--CAGGTGA 8473
QY 381 ----- 381
Db 8474 GAGCTGCATGAGACTGAAGAAAACACTATCAGGCTGAGCCCCACCTAGCTACCA 8533
QY 382 -----ArgAsnGlyIleTyrProLeuMetAsnPheAlaAlaThrAsp 396
Db 8534 CCCCCTGTCTATCCCAAGAAACGGGATCTATCTCACTGATGAACCTTGGCGCTGGC 8593
QY 396 roLeuGlyLeuProArgValIleuAlaProProProGluGluValGlnIysAlaIysThrP 416
Db 8594 CTTGGGGCTTCCCGTGTGTGTGGCCCAACCCCAAGAGAAACCCAAAGCCAAAACTC 8653
QY 416 roThrProGluProPheAspSerGluThrArgValIle----- 429
Db 8654 CAACGCCAAGAACCTTTGACTCGAGAACCAAGAAAGGTAG--TCCCACTCTGTGTACT 8712
QY 429 ----- 429
Db 8713 GTGCTGTCTTAGAGTTGACCAATCAGGAATGGAATGAGATCTTTCTTGAGGA 8772
QY 429 ----- 429
Db 8773 CCGGCTCTCCCATCCCGGTTCTCCCGGAGCTGCAGAGACCAATCCAGCTGACCC 8832
QY 429 ----- 429
Db 8833 CTTTGTGAAGCTGGAACCGGACCGACACATCATTTCCCAAGAGCTTCTTGCTTTAC 8892
QY 429 ----- 429
Db 8893 AGTCTTCCCATCCGTTGATGCTCTTTCAGAGGCTTGGCTACGCTGTCCCC 8952
QY 429 ----- 429
Db 8953 GTCTGATCCAGAGCATTTGCTCCGCTCTCCCTTCTCATAGCTTCTTGTTCTCTAATA 9012
QY 429 ----- 429
Db 9013 CACCTTAATTCATTTGCAGTAGTCTATCATGAGCTGCCAGAGAACTAGTGAATCAG 9072
QY 429 ----- 429
Db 9073 CCTAAGCCCAAGCTCTGTCTCTCTCTGTGGACACTTGTGTTTGTGTACAGTACAC 9132
QY 429 ----- 429
Db 9133 CTTGGGGGACAGGCTTCTTGTGCTTCCAGGTTCTCTTCTGTCTGTCTGTGC 9192
QY 429 ----- 429
Db 9193 CTCCCATCCCTCTCTGGCATTTGATTCATCATTCAGAGACCAACCGAAGCTGTCT 9252
QY 429 ----- 429
Db 9253 TGAAGCTTCCCTAATACTCAGGGCTGTGTGAGTAAGAGTGTCTTTCGACTTCCCTTG 9312
QY 429 ----- 429
Db 9313 GAGACATTATACATCTTGGGTCTGTGTCTCTCTTCTGGGCTCTCTGTAAAGCCGTG 9372
QY 430 -----GlnMetGlnCyAsnLeuGlu 437
Db 9373 CTCTGTGCTAATACACTTCCCTGCCACAGGTGTTCAGATGCACTGCAACTGTGAAA 9432
QY 437 rGSerGluAspIysAlaArgTyrHis----- 445
Db 9433 GAAGGAGGACAAAGCTCGGTGGCAGTGAAGCAAGGGCTGTGTGAGACTAGGGAAGAT 9492
QY 446 -----LeuThrLeu 449

DB	9493	GGGAGCAATCAACTCTGGCGCGCTTGACACAGCCCTTCGTTGAGCTTACTCTGC	9552
OY	449	euLeuValleuGlUaApArIgleuHlSaIglInLeuThrTYraSpLeuLeuProthr	467
Db	9553	TCCTTGTCCTTGAAGACCGGCTACTATCGGACAGTACCTAATGATCTGCCAAG-TAGG	9611
OY	467	-----	467
Db	9612	CTGCTAGTGGGTCAGGTGTGGGCTGAGGCTCAGGACCTAGGCTATGTCCTACTATT	9671
OY	468	-----AppSerAlaGlnAspLeuAlaSerGluLeuValHisT	480
Db	9672	TGCTCTGTCTCTCTCTTACAGCGACAGTCCGACAGACTCGCTCTGAACTAAGTCACTT	9731
OY	480	YTGlyPheLeuHlSglU-----	485
Db	9732	ATGGCTTCTTGACAGA-GATGGCGCGGTGCGCTTGTGGTGTGAAGGTGACGGTCTGC	9790
OY	486	-----AspAraPrgMetIleuVal	491
Db	9791	CGAGTGGCCGGTACTACCATGCTTCCCTACTCATCCATCGAGGATATACAGGACAAAGC	9850
OY	491	euAlaAlaPheLeuGluSerThrPheLeuIleuTYraGlyThrGlnAla	507
Db	9851	TAGACGCTTCTTGAGACACACTTTCTTCAGACTACCGAGGACGCAAGC	9900
RESULT 10	AC110211/C	202844 bp	DNA linear ROD 05-NOV-2004
LOCUS	AC110211		
DEFINITION	Mus musculus chromosome 15, clone RP23-192821, complete sequence.		
ACCESSION	AC110211		
VERSION	AC110211.16	GI:55416090	
KEYWORDS	HTG.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognatha; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 202844)		
TITLE	Birren,B., Nusbaum,C. and Lander,E.		
JOURNAL	Mus musculus chromosome 15, clone RP23-192821		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 202844)		
	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Baetsen,V., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Gellagan,J., Gaidyna,S., Glinde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,M., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kelle,C., Lacroque,K., Lamarez,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McDonald,P., McKernan,K., Meldrum,J., Meneus,L., Minova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicoli,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strausn,K., Subramanian,A., Talamas,J., Teeffaye,S., Theodore,J., Topham,K., Traversman,M., Travis,N., Trigilio,J., Vassiliou,H., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-Feb-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	3 (bases 1 to 202844)		
AUTHORS	Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Archchi,H.M., Barna,N., Baetsen,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,		

```

TITLE
JOURNAL
AUTHORS
REFERENCE
COMMENT
JOURNAL
TITLE
SOURCE
FEATURES

```

Dearlano, K., Diaz, J.S., Dodge, S., Doolley, K., Dorrie, L., Erickson, J., Fero, S., Ferreira, P., Fitzerald, M., Gage, D., Hagopian, D., Gaidyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, R., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, D., Jones, C., Kamat, A., Karatas, A., Kelly, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., McLaren, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menais, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Reikhsang, P., Pierre, N., Rachpak, A., Ramasamy, U., Raymond, C., Reth, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupp, C., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talmas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vasilev, H., Venkatarman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (05-OCT-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 202844)
 Birren, B., Nusbaum, C., Lander, E., Abouelell, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Baetien, V., Bloom, T., Bonslavsky, L., Bouckgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corru, B., Dearellano, K., Diaz, J.S., Dodge, S., Doolley, K., Dorrie, L., Erickson, J., Fero, S., Ferreira, P., Fitzerald, M., Gage, D., Hagopian, D., Gaidyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, R., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelly, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., McLaren, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menais, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhang, P., Pierre, N., Rachpak, A., Ramasamy, U., Raymond, C., Reth, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupp, C., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talmas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vasilev, H., Venkatarman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (05-NOV-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Nov 5, 2004 this sequence version replaced gi:53793775.
 All repeats were identified using RepeatMasker:
 Smit, A.P.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

 Genomic Center
 Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WTBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@broad.mit.edu

 Project Information
 Center project name: L20862
 Center clone name: 192_B_21

 Location/Qualifiers
 1..202844
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="15"
 /map="15"
 /clone="RP23-192B21"
 /clone_lib="RPCI-23 Female Mouse BAC"
 complement(1609...1757)
 /rpt_family="B1_MM"
 7176..7216
 /rpt_family="(CCG)n"
 9059..9095
 /rpt_family="(GA)n"

```
repeat_region      11282..11303
/rpc_family="AT_rich"
repeat_region      11374..11512
/rpc_family="HAL1"
repeat_region      11589..11721
/rpc_family="B1_MM"
repeat_region      11799..11947
/rpc_family="B1F"
repeat_region      12054..12092
/rpc_family="(CA)n"
repeat_region      12349..12472
/rpc_family="B1_MM"
repeat_region      complement(12490..12541)
/rpc_family="B1_MM"
repeat_region      12558..12598
/rpc_family="Lx4"
repeat_region      12652..12978
/rpc_family="Lx4"
unpure             complement(13072..13151)
/note="single clone coverage"
unpure             complement(13142..13146)
/note="30 Qual SINGL region"
repeat_region      13658..13820
/rpc_family="RSINE1"
repeat_region      13829..13867
/rpc_family="(CA)n"
repeat_region      13876..13952
/rpc_family="HAL1"
repeat_region      14121..14197
/rpc_family="B1_MM"
repeat_region      14199..14665
/rpc_family="MT2B"
repeat_region      14826..15002
/rpc_family="MT2B"
repeat_region      14892..15021
/rpc_family="B1_MM"
repeat_region      15022..15043
/rpc_family="AT_rich"
repeat_region      15241..15692
/rpc_family="MT1B3"
repeat_region      15693..15723
/rpc_family="(TG)n"
repeat_region      15724..15776
/rpc_family="MT1B3"
repeat_region      complement(15890..16025)
/rpc_family="L2"
repeat_region      20971..21109
/rpc_family="C-rich"
repeat_region      22357..22540
/rpc_family="B3A"
repeat_region      22726..23017
/rpc_family="L2"
repeat_region      23800..23946
/rpc_family="MT1F2"
repeat_region      24599..24743
/rpc_family="B1_MM"
repeat_region      complement(24761..24909)
/rpc_family="MT1A"
repeat_region      complement(24977..25152)
/rpc_family="B3"
repeat_region      25153..25466
/rpc_family="ORR1A3"
repeat_region      25509..25603
/rpc_family="PB1D10"
repeat_region      complement(25680..25823)
/rpc_family="B1_MM"
repeat_region      25853..26024
/rpc_family="B2_Mm1"
repeat_region      27294..27381
/rpc_family="(GA)n"
repeat_region      complement(27387..27494)
/rpc_family="PB1D10"
repeat_region      27735..27802
```

```
repeat_region      /rpc_family="(GAAAA)n"
complement(28677..28873)
/rpc_family="URR1A"
repeat_region      29722..29903
/rpc_family="L1M2"
repeat_region      30219..30415
/rpc_family="B3"
repeat_region      30422..30447
/rpc_family="AT_rich"

Alignment Scores:
Pred. No.:      1,56e-121      Length:      202844
Score:          1563.50      Matches:      469
Percent Similarity: 32.50%      Conservative: 12
Best Local Similarity: 31.69%      Mismatches: 24
Query Match:      58.56%      Indels:      984
DB:              9      Gaps:      19

US-10-618-941-99 (1-507) x AC110211 (1-202844)

QY      1      MetAla1aProGluProlaProAArgAlaArgLuarGluArgLuar 20
      21043      ATGGCGGCGCCCGGAGCGCGCGCCGAGAGAGGC-----CGGAGCGCGAGCGGAGAGC 20390
QY      21      GluSerGluAerGluSerAerAerAerAerAerAerAerAerAerAer 40
      20989      GAGAGCGAGAGCGAGAGCGAGAGCATCTGAGAGAGAGCCGAGTGGCGCTGGAGAGAGCGG 20330
QY      41      ArgGluGluVal----- 44
      20929      CGGAGAGAGAGT-gggGACGGGTGGCGGTGACCGCGCGTCCGTAAGCATAGATAGAG 20871
QY      44      ----- 44
      20870      CCATGGAGATCCGTGGGGCGGAGCCAGATAGAGGCTCTCCCGAGCGCTGGCGCTTGA 20811
QY      44      ----- 44
      20810      GGCCCTGGATCCTTTGGGCCCAAGCCCTGTGTAGAAACCCCATTCCTTGAACACCGG 20751
QY      44      ----- 44
      20750      TGTGATAGTGGAGACACCTCCAGCGTCTGTGATCCCTACTACCGGCTTGGGGATCTC 20691
QY      44      ----- 44
      20690      CCCAAGCTTAAGACCTGCGTACTCTTGGGCCAGCCCTGGAGACCAACCTTCTCCCC 20631
QY      44      ----- 44
      20630      TGGGTTCTTGATTCACAGCTTACCTCTCAGCCCTCCACTGAGCTGGGAGAGCTAGC 20571
QY      44      ----- 44
      20570      AGCAGATTACAGCGTGGGGGAGAGTTGGGGAGAGCAAGTCCAGACCCCTGGGCTATT 20511
QY      45      ----- 45
      20510      CGTATCAGACACACATTCCTTCCACCTAGGTGAGCAAGGAGCAATGCCCGGATCCAG 20451
QY      54      SerThrPheLeuAlaMetCAspThrGluGluValGluValValLTPAengLLeuHis 73
      20450      AGCAGCTTCTGGGCAAGAGACGAGAGAGGCTGTAGAGTGTGTGAGAGCAACTTAC 20391
QY      74      PheGlyAerArgLysAlaPheAlaAlaHisGlu----- 84
      20390      TTGGGGAGTAGAAGGCTTTGGGGCCATGA-gGTAGGCGCTGTATACCCGCTTGC 20332
QY      85      ----- 85
      20331      CCCATACCGTTGATGGCCCAACCAATCCTTCTCCAGGAGAGATCCAGACCATG 20272
QY      91      PheGluGluLeuValLeuValAerPheProAsnLeValLysLeuHisLysTyrTTPhe 110
```


**Estimated insert size: 212917 - sum-of-contigs
Quality coverage: agarose-FP - N/A
Quality coverage: 6.8 x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 46 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
71517: contig of 71517 bp in length
71518 71537: gap of unknown length
71538 117192: contig of 45655 bp in length
117193 117212: gap of unknown length
117213 133385: contig of 16173 bp in length
133386 133405: gap of unknown length
133406 154137: contig of 20732 bp in length
154138 154157: gap of unknown length
154158 163981: contig of 9824 bp in length
163982 164001: gap of unknown length
164002 169740: contig of 5739 bp in length
169741 169760: gap of unknown length
169761 173809: contig of 4049 bp in length
173810 173829: gap of unknown length
173830 176074: contig of 2245 bp in length
176075 176094: gap of unknown length
176095 178772: contig of 2678 bp in length
178773 178792: gap of unknown length
178793 181465: contig of 2673 bp in length
181466 181485: gap of unknown length
181486 182101: contig of 616 bp in length
182102 182121: gap of unknown length
182122 184693: contig of 2572 bp in length
184694 184713: gap of unknown length
184714 185952: contig of 1239 bp in length
185953 185972: gap of unknown length
185973 188487: contig of 2515 bp in length
188488 188507: gap of unknown length
188508 189224: contig of 727 bp in length
189225 189254: gap of unknown length
189255 190461: contig of 1207 bp in length
190462 190481: gap of unknown length
190482 190657: contig of 186 bp in length
190658 190687: gap of unknown length
190688 192310: contig of 1623 bp in length
192311 192330: gap of unknown length
192331 194183: contig of 1853 bp in length
194184 194203: gap of unknown length
194204 195637: contig of 1434 bp in length
195638 195657: gap of unknown length
195658 196390: contig of 733 bp in length
196391 196410: gap of unknown length
196411 196937: contig of 527 bp in length
196938 196957: gap of unknown length
196958 198075: contig of 1118 bp in length
198076 198095: gap of unknown length
198096 199358: contig of 1263 bp in length
199359 199378: gap of unknown length
199379 200436: contig of 1058 bp in length
200437 200456: gap of unknown length
200457 201048: contig of 592 bp in length
201049 201068: gap of unknown length
201069 201765: contig of 697 bp in length
201766 201785: gap of unknown length
201786 202164: contig of 379 bp in length
202165 202185: gap of unknown length
202186 202997: contig of 813 bp in length
202998 203017: gap of unknown length
203018 204899: contig of 1882 bp in length
204900 204919: gap of unknown length
204920 205365: contig of 446 bp in length

205366 205385: gap of unknown length
205386 205759: contig of 374 bp in length
205760 205779: gap of unknown length
205780 205953: contig of 174 bp in length
205954 205973: gap of unknown length
205974 206484: contig of 511 bp in length
206485 206504: gap of unknown length
206505 206764: contig of 260 bp in length
206765 206784: gap of unknown length
206785 207831: contig of 1047 bp in length
207832 207852: gap of unknown length
207853 208358: contig of 507 bp in length
208359 208378: gap of unknown length
208379 209487: contig of 1109 bp in length
209488 209507: gap of unknown length
209508 210063: contig of 556 bp in length
210064 210083: gap of unknown length
210084 210488: contig of 405 bp in length
210489 210508: gap of unknown length
210509 211550: contig of 1042 bp in length
211551 211570: gap of unknown length
211571 212106: contig of 536 bp in length
212107 212126: gap of unknown length
212127 212454: contig of 328 bp in length
212455 212474: gap of unknown length
212475 212743: contig of 269 bp in length
212744 212763: gap of unknown length
212764 213236: contig of 473 bp in length
213237 213256: gap of unknown length
213257 213817: contig of 561 bp in length.

Location/Qualifiers
1. .213817
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/clone="RP21-208H22"
/sex="male"
1. .71517
/note="assembly_name:Contig205"
71518. .71537
/estimated_length=unknown
71538. 117192
/note="assembly_name:Contig199"
117193. .117212
/estimated_length=unknown
117213. 133385
/note="assembly_name:Contig198"
133386. 133405
/estimated_length=unknown
133406. 154137
/note="assembly_name:Contig197"
154138. 154157
/estimated_length=unknown
154158. 163981
/note="assembly_name:Contig196"
163982. 164001
/estimated_length=unknown
164002. 169740
/note="assembly_name:Contig195"
169741. 169760
/estimated_length=unknown
169761. 173809
/note="assembly_name:Contig194"
173810. 173829
/estimated_length=unknown
173830. 176074
/note="assembly_name:Contig193"
176075. 176094
/estimated_length=unknown
176095. 178772
/note="assembly_name:Contig191"
178773. 178792

QY 255 -----MetAlaVal 257
 Db 14768 TGCTCAGGGGTCTAGCCCATCCCAACCAATTTGTCCTCGACCCAGATGGCTGTA 14827
 QY 258 LeuGluIleGlnThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArg 277
 Db 14828 CTCGAGATCCAGGCCAACCGGGGATACCAAGATCAAGAAAGGCCATGCTCGAGCCAG 14887
 QY 278 HisSerLeuSerAspProAsnMetArg----- 286
 Db 14888 CACTCACTAGTGAACCCCAACATGCGGTGAGACGCTTACCTTGCCCTCGGGGCTGAC 14947
 QY 286 ----- 286
 Db 14948 TGCAGGTGCTTACGCTTACCTTGAAGTGTCTAGGACAGGCCAGAGGGGACGCTGGTG 15007
 QY 286 ----- 286
 Db 15008 GTAAATCTGAGAGGGGACGCTGGTCAATACATCTGAGAGTGTAGCAGAGGACTTAAGGC 15067
 QY 286 ----- 286
 Db 15068 CCACTGAAGAAAGACTCTTCCAGAGCTCAAGACTCCAGAGGGTGAATGAAAGTAG 15127
 QY 286 ----- 286
 Db 15128 GCATACGTGTGGAGCCCAAGGCTTGCTGCTTCAAGTGTCCAGAAATCAAGATCTTGAG 15187
 QY 286 ----- 286
 Db 15188 CTGTAAATGGAGTGAATCTTGTGTCTCATTCATACAGAGAGGCTTTTACATCCCTGAG 15247
 QY 286 ----- 286
 Db 15248 GTGGAGGATGTGTGATTTGACCCGTGGGGTCTAGAGTCTTGAAGACCTCTCTAGAG 15307
 QY 287 -----GluPheIleLeuCy 291
 Db 15308 CGAGACCTTTTCCAGCTATCACCTGCTTTCTTTTACACCTTGAAGAAATGATCTCTC 15367
 QY 291 sCyAlaLeuAlaArgAspProAlaArgAspProSerAlaHisSerLeuLeuPheHisArgVal 311
 Db 15368 CTGGCTGGGCGGGAGACCTGGCCGCGAGACCTCAAGCCCAACCTCTCTTCCACGAGT 15427
 QY 311 lleuPheGluValHisSerLeuLeuLeuAlaAlaHisCyAspHeIleGlnHisGln-T 331
 Db 15428 GCTCTTTGAGGTGCACTGCTGAGCTGTGGCAGCTCACTGCTTCAATCCAGACAGTG 15487
 QY 331 YF----- 331
 Db 15488 TGAGGGGCAAGAGGCCGGGGTGGGTCAAGATCTGGGGTCAAGTATGGAAGGCTTTT 15547
 QY 332 -----LeuMetP 334
 Db 15548 GCAGGACCTTAGTGTGGGTCAATCCCATGAGCATTTGTGTCTTCAACAGACCTCAATC 15607
 QY 334 TGGluAsnValValGluGluValThrLysAlaMetAspLeuHisAlaValLeuAlaGluL 354
 Db 15608 CTGAACAATGTGTAGAGAAAGAACCAAGGCCATGGAATCATATCAGATTTGGCTGAGA 15667
 QY 354 euProArgProArgArgProProlLeuGlnTPrArg----- 365
 Db 15668 TGCCGAGCCCATGAGACCCCAATGCAAGTGGCG-GTGAATAGTGGCCAAACGAGCCCTC 15726
 QY 365 ----- 365
 Db 15727 CCTTCTAGCTGCTTGGCTAGGCTTTGCCCTGACATCTCTCATATGATCTTCCACACAT 15786
 QY 366 -----TyrSerGluValSerPheMetGluLeuAspLysPheLeuGluAspVal----- 381
 Db 15787 CTAGTACTCAGAGGCTCTCTTCTTGGAGCTGGAACAATTTCTAAGAGATG-CAGGTGA 15845
 QY 381 ----- 381

Db 15846 GAGCTGCATGAGACTGAAGAAAAAGCACTATCAGGCTGAGCCCCCACTAGTACCA 15905
 QY 382 -----ArgAsnGlyIleTyrProLeuMetAsnPheAlaAlaThrArgP 396
 Db 15906 CCCCTTGTCTATCCCAAGGAACGGGATCTATCACTGATGAATCTTGGGGTGGCTGGC 15965
 QY 396 ToleuGlyLeuProArgValIleuAlaProProGluGluValGlnLysValAlaLysThrP 416
 Db 15966 CCTTGGGGCTTCCCGTGTGTGGCCCAACCCCAAGAGAACCCAAAGGCCAAACATC 16025
 QY 416 roTHProGluProPheAspSerGluThrArgLysValIle----- 429
 Db 16026 CAACGCCAGAACCTTTGATCTGGAGACCAAGAAAGGTGAG-TCACACTCTGCTGTACT 16084
 QY 429 ----- 429
 Db 16085 GTGTGTCTCTAGAGATTGACCAACAATCAGGGAATGAAATGATGATGCTTCTGGGGA 16144
 QY 429 ----- 429
 Db 16145 CCGGCTTCCCATCCCGTTCCTCCCAAGCTGCAGAGACAAATCCAAGCTGACCC 16204
 QY 429 ----- 429
 Db 16205 CTTTGTGAAGCTGGAACCGACCGACACATCATTTCCCAAGGCTTCTTGTTC 16264
 QY 429 ----- 429
 Db 16265 AGTCTTCCCATCCGGTGTGCAATGCTCTTCCAGAGGCTTTGGCTAGTGTCTGCC 16324
 QY 429 ----- 429
 Db 16325 GTCTGATCCAGAGCAATGTCTCCGGTCTCCCTTCTCAATGCTTCTTCTCAATA 16384
 QY 429 ----- 429
 Db 16385 CACCTTAATTCATTCGACATGATCTACATAGGCTGCCAGAGAACTAGTGAATACTAG 16444
 QY 429 ----- 429
 Db 16445 CCTAAGCCCAAGCTCTGTCTCTCTCTGTGTGACCTTGTCTTCTATACATCAGAC 16504
 QY 429 ----- 429
 Db 16505 CTTGGGGGACAGCGTCTTCTTCTTCCAGGTTCTCTTCTGTCTGTCTGTG 16564
 QY 429 ----- 429
 Db 16565 CTCCCCCATCCCTCTGTGGCATTTGATTCATTCATTCAGACCCAAACGAAAGCTGTCT 16624
 QY 429 ----- 429
 Db 16625 TGAAGCTTCCCTAATCTCAGAGGCTGTGACATGAAGCTGCTTTCCGACTTCCCTTG 16684
 QY 429 ----- 429
 Db 16685 GAGACATTTACATCTTGGGTCTGTCTGTCTCTCTTGGGCTCTCTGTAAGCCGCTG 16744
 QY 430 -----GlnMetGlnLysAsnLeuGluVal 437
 Db 16745 CTCTGTGCTAATACCTTCCCTGCCCAAGGTGTCTCAATGTGACAGTCTGGAAA 16804
 QY 437 TGSerGluAspLysValaArgTTrpHis----- 445
 Db 16805 GAAGGAGAGCAAGCTCGGTGTGCACTGAGACAGAGGCTGTGTTGAGACTAGGAAAT 16864
 QY 446 -----LeuThrLeuL 449
 Db 16865 GGGGACCAATCACTGTGGCGCGCTGACCAAGCCCTTCCGCTGTGAGCTTACTCTGC 16924
 QY 449 euleuValleuGluAspArgLeuHisArgGlnLeuThrTyrAspLeuLeuProThr----- 467

```

Db      16925 TCTGTGCTTGTAGAGACCGGCTACATCGGAGCTGACCTATGATCTGCTCCGAAAG-TAGG 16983
Oy      467 -----
Db      16984 CTGTAGTGGGGGTGAGTGGGCTGAGGCTGAGGCACTGGCTAGTGCATGACTCATT 17043
Oy      468 -----
Db      17044 TGTCTGTGTTCTCTCTCTTCCAGCGGACAGTCCCGACGACTGCTGCTGAACTAGTGCAATT 17103
Oy      480 YTGlyPheLeuH1ag1u-----
Db      17104 ATGGCTTCTGTGACGCA-GATGGGCGGGTGGCGCTTGGGTTGGAAAGGTGACAGGCTTGC 17162
Oy      486 -----
Db      17163 CGAGCTGCCGGTACTACATGCTTCCCTACTCATCATGCGAGGATGACAGCAAGC 17222
Oy      491 eua1a1aPheLeuH1ag1uSerThrPheLeuYb1YrYrG1YThrG1na1a 507
Db      17223 TAGCAGCTTCTTGGAGACACTTCTTCAAGTACCGAGGACGACGCG 17272

RESULT 12
AC132640/c 194709 bp DNA linear HTG 15-NOV-2002
LOCUS Rattus norvegicus clone CH230-339B14, WORKING DRAFT SEQUENCE.
ACCESSION AC132640.3 GI:25007393
VERSION HTG, PHASE2; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Rattus.
1 (bases 1 to 194709)
Muny, D. Marie., Metzger, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,
Ayala-Bech, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Blyant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., P.Souza, L.,
Dayla, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Derramo, C., Ding, Y., Dinh, H., Ditya, K.,
Diexper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgiev, S., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Guarinate, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A.,
Jacksen, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Joliver, A.,
Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, J., Kovar, C.,
Kowals, C., Kraft, C. T., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshewa, L., Louisedge, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhiley, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milibavljovic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L.,
Narkervic, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwokilemeh, O., Okunonu, G., Olarunpasegun, A., Pal, S., Parks, K.,
Pacternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plompe, F., Polindexter, A., Popovic, D., Primm, E., Pu, L. L.,
Puzo, B., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,

```

```

TITLE JOURNAL
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 194709)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submision
JOURNAL Submitted (02-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 194709)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submision
JOURNAL Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
COMMENT On Nov 15, 2002 this sequence version replaced gi:2309481.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center Project name: KBQF
Center Clone name: CH230-339B14
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 188047 bases at least Q40
Consensus quality: 189447 bases at least Q30
Consensus quality: 190412 bases at least Q20
Estimated insert size: 195068; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Gendank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 194709: contig of 194709 bp in length.
Location/Qualifiers
1. 194709
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-339B14"
1. 3065
misc_feature

```

[illegible]

Db	54643	TTTGGCCCTATACCGTTGAATTGCCCCCAGACCAATCTTCTTCCAGAGAAAGATCCAG		54584
Qy	89	ThrValPhegluglnLeuValLeuValAspHisProAsnIleValIysLeuHisIysTy		108
Db	54583	ACCATGTTTGGACAGCTGGCGCTGTGTGGACACCCCAACATCTCTCAAGCTGCACAGTAT		54522
Qy	109	TrpLeuAspThrSerGluAla-----		115
Db	54523	TGGCTGATGCTCTTGAGGC-CCGCGGAGGTGAGCTAGGGCGGGTGAAGGCGGAGT		54465
Qy	116	-----Cys		116
Db	54464	GAAAGCCACAGGCTGGGTGGGACAGATGGGAGCCACCCGCAACGGCGAGACCCGCTTCG		54405
Qy	117	AlaArgValIlePheIleThrGluTyValSerSerGlySerLeuIysGlnPheLeuIys		136
Db	54404	CCGAGGCTATCTTCATCACAAGAAATAGTGTCTGTGGACGCTCAAGAGTTCTCTCAA		54344
Qy	137	LysThrIysIysAsnHisIysValIleMetAsnAlaArg-----		148
Db	54344	AAAGCCAAAGAAACCAAGAGCCATGAAAGCCGGGTATGGGAGTGGGTGGGAGCTG		54288
Qy	148	-----		148
Db	54284	GAGGAGTCCAGACAGAACTTAGGGAAGTATGGGCGGGGAGCTCCGAGAGGCTCGGT		54225
Qy	148	-----		148
Db	54224	TCAGAGTGGGGTATAGCGGCGGGCTCCGAGGCTCAGACGCTCTTGTTGCCACCCGAC		54165
Qy	149	-----AlaTrpIysArgTrpCysIleGlnIleLeuSerAla--		160
Db	54164	GGAATCGTGGCTCGGCTGCGACGAGCTCGAAAGCGCTGTGTATGGCAATCTGTCTGCACT		54105
Qy	160	-----		160
Db	54104	CAGGTAGAGCTGGCTTGGCCACGACCTTGCCCCCTGCTTCTTCACTTCTTCTCC		54045
Qy	161	-----Ileu-SerPheLeuHisAla		166
Db	54044	GGTCCGCTGACCCCGGTTTTCTCTGCTCCCTCTCTCTATGCTGTCAAGTCTTTTGGACG		53985
Qy	166	IaCysSerProProlIleIleHisGlyAsnLeuThrSerAspThrIlePheIleGlnHisAla		186
Db	53984	CTGCAAGTCCCCCATCATCATCAAGAACTTGAACAGGACACAGATCTTCAATTCAACACAA		53925
Qy	186	snGlyLeuIleIysIleGlySer-----		193
Db	53924	ATGACCTCATCAAGATCGGCTGTGTGTGCGAGCGGGCGGGGACAGGGGAGGGGT		53865
Qy	193	-----		193
Db	53864	GGGAAACAGCGGAGCCACAGACAGAGACTGCCAGTCTCTGGCCCCACTGACGCCCT		53805
Qy	194	-----ValTrpHisArgIlePheSerAsn-----		201
Db	53804	CTCTTCTCTGTCGCTCGGCGCGGTGAGGTGTGGTACCGATCTTCTCCATGTGTGTGCG		53745
Qy	202	-----AlaLeuArgProProThr		207
Db	53744	GGTCTCGGGGTGGAGCTGTGAGAAACTGTGTCTGCTCAGCATCAACATTTTCTTAA		53685
Qy	208	AlaLeuProAspAspIleuArgSerProIleArgAlaGluArgIuglnLeuArgAsnLeu		227
Db	53684	GCACTTCTTGACATCTCGAGACCTCTATTAAGAACTGAACGGGAGAGCTTCGAAACCTG		53625
Qy	228	HisPhePheProGluIleTyGly-----		235
Db	53624	CACCTTTTCCCAACCAAGTATGGCGGTAGTCTTAAAGCCGCTTCCATCTGACTGTGCTG		53565
Qy	236	-----GluValAlaAspGlu		240

Db	53564	GCAATCATGAATGCTCCATCAATCACTCTGCTTTGTTTTCATTTCCATTTCAAGAGTCAAGATGG	53505
OY	240	YThAlaValAapIlePheSerPheGlyMeCyAlaLeuGlu-----	254
Db	53504	GACTGCTGTGGACATCTTCTCTTTGGGAGTGTGGACCTGGA--GGAATCTTTCATCTTCC	53446
OY	254	-----	254
Db	53445	TCATAGCCTCTGCTTTGGCTCTACTCCATCTCAACAATTCCAAATTGTGTCCTGCACCCA	53386
OY	255	-MeAlValLeuGluIleGlnThrAengIyaAPThrArgValThrGluGluAlaIleAl	274
Db	53385	GATGGCGGCTTTGGAGATTCACAGCCAGGGGGAATCCCGAGTCAAGAAAGATTATTC	53326
OY	274	aArgAlaArgHisSerLeuSerAapProAanMetArg-----	286
Db	53325	TGGGGCCAGGACCTCGCTGAGTGAACCCAAATGCGGGTGAAGACGTTAAACCCCTTAGG	53286
OY	286	-----	286
Db	53365	ACTGCTGACAGTCTTGAGGCTTGCCCTTGAAGCCCTTAGTAGGCTAGAGGTAGTCT	53206
OY	286	-----	286
Db	53205	GGTACATACATTTGAAGTGTTAGCATCTGAAGCGTAGCTGGCGGACTGAGGTCACTTC	53146
OY	286	-----	286
Db	53145	AAAGAAAGACTCTTCTAGAGACTCAGGGAATTTCAAGGGGTGAATGAAGTAGGCAATG	53086
OY	286	-----	286
Db	53085	GGTGTTTGAGCCCAAGGCTTCTCTGCTCTCAAGTTTACAGAATCAGATCTGAGTCT	53026
OY	286	-----	286
Db	53025	AATGGATGCCCTTCTCAAGGACGCAATGAGAGACATCATTTGTGAGGTGGAGAGATCC	52966
OY	286	-----	286
Db	52965	AAATGATTGACCCCGTGGGCTTAGGGCTTGAAAGATCTCTCTTACAGGAGATTTTCC	52906
OY	287	-----GluPheIleLeuCySValLeuAlaArgA	296
Db	52905	ACCACCACTACCTTCTCTCCCAACACTAGGAATTCATCTCTGCTGCGCCGGG	52846
OY	296	gPrProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGluValH	316
Db	52845	ACCTGCCCCCGGACCTTCAGGCCACCACTCTCTTCCACCGAGTGCCTTTGAGGTGC	52786
OY	316	IsSerLeuValLeuLeuAlaAlaHisCyAPheIleGlnHisGln--Tyr-----	331
Db	52785	ACTGCTGAAGCTCTGAGCAACCACCTGTTTCATCCAGCACAGTGTGAGGGGAGAAAG	52726
OY	331	-----	331
Db	52725	GCCGGGTGACCAAGATCTGGGGTCAGATCATGGAAGGGTCTTTGCAGAACCTTAGG	52666
OY	332	-----LeuMetProGluAanValVal	338
Db	52665	TTGGGTCCATCCCAATGAGGCAATTTCTGTGTTCACAGACTTCATGCTAGATGTGGTA	52606
OY	339	GluGluIysThrIysAlaIleMetAapLeuHisAlaValLeuAlaGluLeuProArgProArg	358
Db	52605	GAGGAAAAGACCAAGGCCATGGAACCTTCATGCAGTGTGGCTAGATGCCCCACGCCCAT	52546
OY	359	ArgProIleuGlnIntPrArg-----	365
Db	52545	GGAACCCCAATGCAGTGGCG--GTGAGTGTGGCAAAATGAGGCTTCCCTCTTAACGTCT	52487
OY	365	-----	365
Db	52486	GGCTAGGCTTTGCTTAGTGCCTTATATGCTTATATGACATCTTCCATACATCCAGG	52427

QY	366	TySergIuValSerPheMetGluLeuAspIysPheLeuGluAspVal	-----	381
Db	52426	TACTCAGAGGTTTCCTTGTGAGCTGCACAAATTCCTAGAGGATGTCAAGTAGAGCTG	-----	52367
QY	381	-----	-----	381
Db	52366	GCATGACAGTGAAGAAAAAACTATCAGACCAGAGTCCCCACCTACCTACCCCTGTCT	52307	
QY	382	-----ArgAnGlyIleIyrProLeuMetAsnPheAlaIaIaThrArgProLeuGlyLeu	399	
Db	52306	ATCTCCAGAGATGGCATCTATTCACATCATGAATTTTGGCTGTGCTGGCCCTTGGGACTT	52247	
QY	400	ProhArgIleuAlaProProProGluGluValGlnIleAlaIleIysThrProThrProGlu	419	
Db	52246	CCTGTGTGTGTGGCCCCACCCCAAGAGAGGCCCAAAAGCCAAATCTCCAACTCCCAAGAA	52187	
QY	420	ProPheAspSerGluThrArgIys-----	427	
Db	52186	CCCTTTGACTCCGAGACACAGAAAGGTGAATCCCACTCTGTGTACTATGTGTGTTTCTC	52127	
QY	427	-----	427	
Db	52126	AGAGTTAACCAACAACAGGGAATGGTGCCTCTGGGACCTGTCTGTGCATCCCAATT	52067	
QY	427	-----	427	
Db	52066	CTCCCCAGAGCTTCCAGAGACCAAAATCCAAGCTGACATTTTTCCCCAAGGCTTCTGC	52007	
QY	427	-----	427	
Db	52006	TTTTACAGTCTTCCCCATCCCATGTCGAGTCTCTTCATAGAGCTTTTGGCACCTCCA	51947	
QY	427	-----	427	
Db	51946	TTCCCCATCTGATCCAGAGCCATCACTGTATCTCCCCATCTAGTTCATCTCTCC	51887	
QY	427	-----	427	
Db	51886	TCAGCACACACTAAATTCAAAATGCAGTGGTCTACATGTCTGSAATCCAGCTGCACAGA	51827	
QY	427	-----	427	
Db	51826	GAGCTAGTGAATAATCAGCCTCAGGGGGTGTGGGATTTAGCTCAGTGAAGAGCGCTTGCC	51767	
QY	427	-----	427	
Db	51766	TAGAAACCTGAAGGCCCTGGGTTGGGTCCCCAGCTCCGAAAAAAGAACCAAAAAAAA	51707	
QY	427	-----	427	
Db	51706	AAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAATCAGCCTCAGTCCCTGT	51647	
QY	427	-----	427	
Db	51646	CTGTCTCTGTGGCACCACCATTTTCTGTACTGTTCGACCTGAGACACAGACTTCTCT	51587	
QY	427	-----	427	
Db	51586	TGCTTTCACAGATTCTCTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT	51527	
QY	427	-----	427	
Db	51526	GTCCTGTCTCTCTGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTG	51467	
QY	427	-----	427	
Db	51466	CATTGTATTATCATTCAGTACCAACAAGAGCTACTACTGTGAAACCTTCCCTAAT	51407	
QY	427	-----	427	
Db	51406	ACCTCAGAGATGTGACGTAAAGATGTGTTTCCGACTTCCCTTGAGATGTCTACATT	51347	

Accession	Version	Keywords	Organism	Reference	Authors
AB011518.4	1	HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.	Rattus norvegicus (Norway rat)	1 (bases 1 to 248568)	Muzny, D., White, M., Metzker, M., Lee, A., Brown, M., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, B., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benham, F., Bernal, K., Blair, J., Blankensburg, K., Blythe, P., Brown, M., Bryant, N., Buhay, C., Burck, P., Burrell, K., Caldeira, E., Cardenas, V., Carter, K., Cavazos, I., Cebasa, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyne, M., Cree, A., D'Souza, L., Davis, C., Davy, Carroll, L., De Anda, C., Dederich, D., Delgado, O., Deng, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotte, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabel, A., Gante, R., Garcia, A., Garner, T., Garza, M., George, G., Geor, K., Gill, R., Gridley, M., Guert, W., Guetara, W., Gunaratne, P., Haaland, M., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Hawak, P., Hayes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,

JOURNAL TITLE
AUTHORS
REFERENCE TITLE
JOURNAL

COMMENT

Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 6, 2002 this sequence version replaced gi:23267220.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly ('a' contig-scaffold'). Within each contig-scaffold, individual sequencing contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GSCY
Center clone name: CH230-S2N10

Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 23402 bases at least Q40
Consensus quality: 233092 bases at least Q30
Consensus quality: 234313 bases at least Q20
Estimated insert size: 237322; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

Db 86874 CAGGTGAGACCTGCTTGCCACGCACTTGCCCTGCTCTTCACTTCTTCC 86933
 Qy 161 -----Leu-SerPheLeuHisA 166
 Db 86934 GGTGCGCTGACCCGGTTTTCCTGCTCCCTTCTCTATGCTGCAAGTTTTCAGG 86993
 Qy 166 1ACySerProProIleIleHisGlyAsnLeuThrSerAspThrIlePheIleGlnHisA 186
 Db 86994 CCGCAGTCCCCCATCATCATCAGCAAGAACTGACGACGACGACGACTTCTTCACTTCCACACA 87053
 Qy 186 ArgGlyLeuIleGlySer----- 193
 Db 87054 ATGGCTCATCAAGATCGCTGTGTGTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 87113
 Qy 193 ----- 193
 Db 87114 GGGGAACAGCGGAGCCGACAGACAGACTGCCAGTCTCTGCGCCCACTGACCGCT 87173
 Qy 194 -----ValTrpHisArgIlePheSerAsn----- 201
 Db 87174 CTCTTCTCTGCTGCGCGCGCTTGCAGTGTGTGTACCGCATCTTCTCCAAATGCTGTGG 87233
 Qy 202 -----AlaLeuArgProProIle 207
 Db 87234 GGTCTCGGGGTGTGGAGCTGTGAGAAACTGTCTGCTCAAGCATCACTTTTCTTA 87293
 Qy 208 AlaLeuProAspAspLeuArgSerProIleArgAlaGluArgGluLeuArgAsnLeu 227
 Db 87294 GCACTTCGACGATCTCGGAGCCCTTATTAAGAACTGAACGGAGAGAGCTCGAAGACCTG 87353
 Qy 228 HisPhePheProGluTyrGly----- 235
 Db 87354 CACTTTTCCACACAGATATGCGGTGAGTCTTAAGCCGCTTCCATCTGACTGCTG 87413
 Qy 236 -----GluValAlaAspGlu 240
 Db 87414 GCACATCAGATATGTCATCACTGCTGCTGTTTCCATTTCCAGAGTCAAGATG 87473
 Qy 240 YhrIleValAlaPhePheSerPheGlyMetCysAlaLeuGlu----- 254
 Db 87474 GACTGTGTGACATCTTCTCTTGGAGTGTGTGACATGGA--GGTACTGTTCATCTGCC 87532
 Qy 254 ----- 254
 Db 87533 TCATAGCTCTGCTTGGCTTACTCATCTCACTTCCATTTGTCCTTCCACCA 87592
 Qy 255 -MetAlaValLeuGluIleGlnThrAsnGlyAspThrArgValThrGluAlaIleAl 274
 Db 87593 GATGGCGGTCTTGAGATCCAGGCCAAGGGGATACCAGATCAGAAAGAAAGCTATTCC 87652
 Qy 274 ArgAlaArgHisSerLeuSerAspProAsnMetArg----- 286
 Db 87653 TGGGGCAGGCACTCGCTAGTGATCCCAACATGCGGTGAGCAGCTTAACCCCTCTAG 87712
 Qy 286 ----- 286
 Db 87713 ACTGCTGAGGCTTGGGCTTGCCCTTGAAGCCCTTAGAGTGAAGGTAGTCT 87772
 Qy 286 ----- 286
 Db 87773 GGTCAATACATTTGAAGTGTAGATCTGAAGAGCTAGCTGCGGAGCTGAGCTC 87832
 Qy 286 ----- 286
 Db 87833 AAAGAGAAGCTTTAGAGCTCAGGAATTTCAAGAGGTGAATGAAGTAGGCGATG 87892
 Qy 286 ----- 286
 Db 87893 GGTGTTTGAAGCCCAAGGCTTCTGCTCCTCAAGTTTACAGATCAGATCTGAGCTGCT 87952
 Qy 286 ----- 286
 Db 87953 AATGGAGTCTTCTTCAGGCAAGGATGATGAGACATCATTTGTGAGGTGAGAGATGC 88012

Qy 286 ----- 286
 Db 88013 AATGATTGACCCGTGGGTTAGGCTTTGAAGAATCTCTTAGAGGAGATTTTCC 88072
 Qy 287 -----GluPheIleLeuCysCysLeuAlaArgA 296
 Db 88073 ACCACCACTACTTCTTCCCCACAACTAGAAATTCATCTCTCTCTGCGCGG 88132
 Qy 296 ArgProAlaArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGluValH 316
 Db 88133 ACCCTGCGCGGACCTCAGCCCAACACTCTCTCTTCCACGAGTGTCTTGGAGTGC 88192
 Qy 316 HisSerLeuValLeuLeuAlaHisCysPheIleGlnHisGln-Tyr----- 331
 Db 88193 ACTGCTGAAGTGTGCTGGAGCCCACTGTTCATCCAGACAGTGTGAGGCGAGAAAG 88252
 Qy 331 ----- 331
 Db 88253 GCGGGGTGAGACGAGATCTGGGTCAGATCATGGAAGGCTTTGACGACCTTAG 88312
 Qy 332 -----LeuMetProGluAsnVal 338
 Db 88313 TTGGTCAATCCCATGCGCATTTCTGTGTTCACAGACTCATGCTGAGATGTGTA 88372
 Qy 339 GluGluValThrIleValAlaMetAspLeuHisAlaValLeuAlaGluLeuProArgProArg 358
 Db 88373 GAGGAAAGACCAAGGCCATGTGACCTCATGTGATGTGGCTGAGATGTCCGAGCCCAT 88432
 Qy 359 ArgProProLeuGlnTyrArg----- 365
 Db 88433 GACCCCCCAATCATGTTGGG--GTGAGTGTGCAATGAGCTTCCCTCTTAATCTGCTT 88491
 Qy 365 ----- 365
 Db 88492 GGCTAGGCTTGGCTTGAATGCTTATGTCTTATGACACCTTCCATACCATCCAG 88551
 Qy 366 TyrSerGluValSerPheMetGluLeuAspIlePheLeuGluAspVal----- 381
 Db 88552 TACTCAGAGTTCCTCTCTTGAGCTGACAAATTCCTAGAGATGTGAGTGAAGCTG 88611
 Qy 381 ----- 381
 Db 88612 GCATGACGTGAAGAAAAACATATCAGACCCGAGTCCCACTTACCCCTGTCT 88671
 Qy 382 -----ArgAsnGlyIleTyrProLeuMetAsnPheAlaAlaThrArgProLeuGlyLeu 399
 Db 88672 ATCTCCAGGAATGCACTATCATCTCATGATTTGCTGTGCTCGGCTTGGGACTT 88731
 Qy 400 ProArgValLeuAlaProProProGluGluValGlnValAlaValThrProThrProGlu 419
 Db 88732 CCTCGTGTGTGGCCCCACCCCAAGAGAGGCCCAAAAGGCCAAATCCCAACCCACAGA 88791
 Qy 420 ProPheAspSerGluThrArgGlyS----- 427
 Db 88792 CCTTTGACTGGAACAGAGAGAGGTGATGCCAACTCTGTGTACTATGCTGTTTCTC 88851
 Qy 427 ----- 427
 Db 88852 AGAGTTAACACACAGGGAATGGGTGCTGTGGGACCTGTCTGTGTCATCCCAATT 88911
 Qy 427 ----- 427
 Db 88912 CTCGCCAGGCTTCCAGAGCAACAAATCCAGTGCATTTTTCACAGAGCTTCTTGC 88971
 Qy 427 ----- 427
 Db 88972 TTTTACAGTCTTCCCATCCCATTTGCGGTGCTCTTCCATAGGCTTTTGGCACCTCCA 89031
 Qy 427 ----- 427
 Db 89032 TTCCCCATGTGATCCAGAGCATCATCTGATCTCCCAATCCAGTTCATCTTCC 89091


```

Oy 427 ----- 427
Db 89092 TCAGACACACTAAATTCAATGCCAGTCGTCTACATGTCCTGGATCCAGCTGCACAGA 89151
Oy 427 ----- 427
Db 89152 GAGCTAGTAAATCAGCCTCAGGGGGTTGGGGATTTAGTCACTGAGTGAAGCGCTTGC 89211
Oy 427 ----- 427
Db 89212 TAGAAGCGTAAGCCCTGGGTGGGTCCAGCTCCGAAAAAAGACCAAAAAA 89271
Oy 427 ----- 427
Db 89272 AAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAATCAGCTCAGTCCCTGT 89331
Oy 427 ----- 427
Db 89332 COTGCTCTGGTGGACCCCAATTTTCTGTACTGTGTGACCCCTGAGACAGCATTTCT 89391
Oy 427 ----- 427
Db 89392 TGCTTCCACAGTCTCTTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTCTT 89451
Oy 427 ----- 427
Db 89452 GTCTCTGCTCTCTCTGTGTGTCTGTCTGTCTCTCTCTCTCTCTCTCTCTCTGG 89511
Oy 427 ----- 427
Db 89512 CATTGTATTCAATCATTAGTACCACAGAAAGGCTACTACTTGGAGCCTTCCCTAAT 89571
Oy 427 ----- 427
Db 89572 ACTCAGGAGATGTCAGTAAGATGTGTTTCCGACTTCCCTGAGATGCTACAT 89631
Oy 427 ----- 427
Db 89632 TTTTGGGTCTGCTGATCTCTTGGGCTCCCTTTTAAATTAATCTCTGGCTAATAC 89691
Oy 428 ----- 442
Db 89692 ACTTCCCTGCACAGTGTGTCAATGATGACCTTACTCTGCTTTGGTGTGAGGA 89751
Oy 442 AATGTPhis----- 445
Db 89752 TCGGTGCACTGAGCAAGGCGCTGGGTGGGACTAGAGAAAGACGAGCAATCACTC 89811
Oy 446 ----- 454
Db 89812 TGGCAGCCGCTGACACAGCCCTTCCGTCGACAGCTTACTCTGCTTTGGTGTGAGGA 89871
Oy 454 PARgleuH1BARglnleuH1rTyAsp1leu1euPro----- 466
Db 89872 CCGGTATATGGGAGCTGACCTATGATCTGTCTCCAGTAGGAGCTGTAAGGGGATCAGT 89931
Oy 466 ----- 466
Db 89932 GGTGGGCTGGGCTCAGGACCGGCTAGTGACCAACTCACTTGTCTGTCTCTCTT 89991
Oy 467 -Thr-AspSer1aGlnaPhleu1aSer1aGlnaVal1aTyrg1aPhleuH1aGln- 485
Db 89992 CAGCGGACAGTCCAGGACCTGCGCGGCAACTAGTCACTATGGCTTCCGACGAGG 90051
Oy 485 ----- 485
Db 90052 TGGGGGGGTGGCGCTCGGGTTTGGAAAGGTGACGTCTGCCAGCTGCCAGACTACCGT 90111
Oy 486 ----- 497
Db 90112 GCTTCCCAACCCCTTCACACTGAGATGACCGGACAAAGCTAGAGCTTCTTGAGAGCC 90171
Oy 498 ThirPhleu1aTyTyArg1aTyThirGln1a 507

```

```

Db 90172 ACTTTCTCAAGTACCTGGGACACAGCG 90201
RESULT 14
BC074684
LOCUS
DEFINITION
Xenopus tropicalis nuclear receptor binding protein, mRNA (cdna
clone MGC:69507 IMAGE:5336424), complete cds.
ACCESSION
BC074684
VERSION
BC074684.1 GI:49250535
KEYWORDS
MGC.
ORGANISM
Xenopus tropicalis (Silurana tropicalis)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE
1 (bases 1 to 3046)
Strausberg,R.L., Pelngold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buettner,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schecter,T.B., Brownstein,M.J., Udell,T.B., Tohilyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mulianby,S.J., Boask,S.A., McEwan,P.J.,
McKernan,K.U., Malek,U.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Bakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E.,
Sutcher,A., Schein,J.E., Jones,S.J. and Marz,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3046)
Klein,S. and Gerhardt,D.S.
Direct Submission
Submitted (23-JUN-2004) National Institutes of Health, Xenopus Gene
Collection (XGC) National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: DOE Joint Genome Institute, Walnut Creek,
California
Web site: www.jgi.doe.gov
REMARK
COMMENT
FEATURES
Source
1..3046
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:69507 IMAGE:5336424"
/tissue_type="Embryo, Silurana tropicalis, tailbud (stages
20-27)"
/clone_id="NICHD_XGC_Emb7"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Series: IRAX Plate: 130 Row: C Column: 1

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

Location/Qualifiers

gene
1. .3046
/gene="nrbp-prov"
/note="synonym: MG669507"
/db_xref="GeneID:448149"
128. .1708
/gene="nrbp-prov"
/product="nuclear receptor binding protein"
/protein_id="AAH74684.1"
/db_xref="GI:49250536"
/db_xref="GeneID:448149"
/translation="MSEGGKTA PSTGDPKMSGVMSPPVAPAPPOEESESE
DSEBIEIPCGRMQKREBNQNPVGDIAVLADTEGVVWNEVQPSRRKFK
MOBEKVAVDNLIQLEHNIYKFKYAVDVKNRAVLPTEWSSGLKQKFKTK
KHKTKNEKAMKMCOTLISLVLSQDPRIHGNLTCDTPIQNGIKIGSVAPD
TNNHYKTRBEGKNIHFPAPREYERTNTTAVDYSFGKCLMVALEIQQGSSSY
VQEDATNNAIQPLEDPQRFQKCLDTPSKRPARELHPALEVPSLKILAAHC
IVGHQIMIPENALEMTKNLDMVALEISHDREGRVMI PQSPALIEDKLEVDN
GIYPLTAFVPRPOQPOEVVKS PVPVPKPTPPAPAVETKVVQOMQCSVDG
AKHILTLKLEDKLNRLSCDLLPNEINIQLAELVQGVSESDQIRLACLEAF
SKFYFARNQSMVAIVTSS"

ORIGIN

Alignment Scores:
Pred. No.: 5, 87e-122 Length: 3046
Score: 1541.00 Matches: 306
Percent Similarity: 72.99% Conservative: 67
Best Local Similarity: 59.88% Mismatch: 102
Query Match: 57.72% Indels: 36
Gaps: 7

US-10-618-941-99 (1-507) x BC074684 (1-3046)

QY 1 MetAlaLeuProGluProAlaProAlaArgGluArgGluArgGluArg 20
DB 209 GTTAGTGGCCCC--CCTGCACCA-----CCACAGAGAGAGAAAGAG 250

QY 21 GluSerGluAspGluSerAspIleLeuGluGluSerProGluArgTrpGlnVal 40
DB 251 GAGAGTGAAGATGAGTCTGATCTTAAGAGAGTCTTGGCGGGCTGGCGAAAGAA 310

QY 41 ArgGluGlnValAsnGlnGluAsnMetProGluLeuGlnSerThrPheLeuAlaMet 60
DB 311 CCGAGGAGAGTGAACCAAGCAATGCTGCTGCATTGACAGTCCCTTGCATGAGAC 370

QY 61 ThrGluGlnGluValGluValValTPraangluLeuHisPheGluAspArgValaPhe 80
DB 371 ACAGAGAGAGAGTGGAGGTGCTCTGGAACGAGGTGCTCTGAAAGAAAGAACTTC 430

QY 81 AlaAlaHisGluGluValLeuGlnThrValPheGluGlnLeuValLeuValAspPro 100
DB 431 AAAATGCAAGAGAGAAAGTGAAGCCGTTGTTGAATCTTATTCAGCTGAGCATCTT 490

QY 101 AsnIleValIleLeuHisValSerTrpLeuAspThrSerGluAlaCysAlaArgValIle 120
DB 491 AACATTGTCAAGTCCACAAAGTACTGGCGCATGTGAAGAGAAACCGAGCCGGGTCTATC 550

QY 121 PheIleThrGluTrpValSerSerGluSerLeuValGlnPheLeuValSerTrpVal 140
DB 551 TTCATCTACAGAGTACTGTCTCTGGGAGTCTCAAGCAGTCTCTCAAGAAAGCAAGAA 610

QY 141 AsnHisValLeuMetSerAsnAlaArgAlaTrpValArgTrpCysThrGlnIleLeuSerAla 160
DB 611 AACCAAAAACCATGATGAATAAAGCCTGGAAGATGTGACGAGATCTCTCCGCT 670

QY 161 LeuSerPheLeuHisAlaCysSerProPheIleHisGluAsnLeuThrSerAspThr 180
DB 671 CTCAGCTACTGCACTCTGTGACCCCCCAATATCAAGAAACCTTGACGTGCGACAA 730

QY 181 IlePheIleGlnHisAsnGluLeuValSerValIleGluSerValTrpHisArgIlePheSer 200
DB 731 ATTTTATTCAGACAAACGAGTCTATCAAGATTTGATCTGTT----- 772

QY 201 AsnAlaLeuArgProProthralLeuProAspAspLeuArgSerProIleArgAlaGlu 220
DB 773 -----GCGCCGAGACCACTTAACACACATGAAAGACCTTC 808

QY 221 ArgGluGluLeuArgAsnLeuHisPhePheProGluTrpGluValAlaAsp--- 239
DB 809 CCGGAGGAAACAGAAAGACCTGACCTTCTTGGCTCCGAGTACGAGAACTAATATG 868

QY 240 GlyThrAlaValAspIlePheSerPheGluMetCysAlaLeuGluMetAlaValLeuGlu 259
DB 869 ACAACTGCGGTGACATTTACTCTCTTGGGATGTGTGCCCTGAGATGCGGTGCTGAG 928

QY 260 IleGlnThrAsnGluValAspHisArg---ValThrGluGluAlaIleAlaArgAlaAsnHis 278
DB 929 ATCTAAGAAACGAGGAGTCTCTGATGTGTTCCCGAGAGGCGATCAATATGCAATACAG 988

QY 279 SerLeuSerAspProAsnMetArgGluPheIleLeuCysCysLeuAlaArgAspProAla 298
DB 989 TTCCTGGAAGATCCGTCGAGAGGAGTTATTCAGAAAGTGCCTTGAGACCGACCCACAG 1048

QY 299 ArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGluValHisSerLeu 318
DB 1049 AAGCGCCCAACAGCCCGGAGTTACTCTTCCACCCCGCTGTTGAGGTGCATCCCTG 1108

QY 319 LysLeuLeuAlaAlaHisCysPheIleGlnHisGlnTrpLeuMetProGluAsnValAla 338
DB 1109 AAATTACTGCTGCGGACATGATTTGGGTACCAACATATGATACAAATAAACCGCTTA 1168

QY 339 GluGluLysThrLysAlaAspLeuHisAlaValAlaGluLeuProArgProArg 358
DB 1169 GAAAGAAATGACCAAAACCTGACATGAGCCGCTCTTGGGAGATCTCGACGAGAC 1228

QY 359 ArgProProLeuGlnTrpArgTrpSerGluValSerPheMetGluLeuAspLysPheLeu 378
DB 1229 CCGGAGGAGCGGAGTGAATATTTCTCCAGTCCCGACGCTGAGCTGATTAAGTCTCG 1288

QY 379 GluAspValArgAsnGluLysTrpProLeuMetAsnPheAlaIleThrArgProLeuGly 398
DB 1289 GAAAGATGTGAAGACGCAATTAACCCCTGACTGCAATTTGGAGTGGCCGCGCT----- 1342

QY 399 LeuProArgValLeuAlaProProProGluGluValGlnLysAla----- 413
DB 1343 -----CAGAACCCCAACAGAGGTGTGTAAATCCCTGTGTTCCCCCA 1387

QY 414 -----LysThrProThrProGluProPheAspSerGluThrArgLysValIleGlnMet 431
DB 1388 TCTGTAAGACCCCTTACCCCAAGCCAGCCGAGGTGAGAACCCGCAAGTGTGAGAGT 1447

QY 432 GlnCysAsnLeuGluArgSerGluAspLysValAlaGTrpHisLeuThrLeuLeuVal 451
DB 1448 CAGTCCAACTATGAGTCTGTGATGAGGCGCCCAAGCATCACTGATTTGCTTAAG 1507

QY 452 LeuGluAspArgLeuHisArgGluLeuThrTrpAspLeuLeuProThrAspSerAlaGln 471
DB 1508 CTGAGAGTAACTGAACAGACACTGAGCTGTGACCTGCTCCCAATGAGAACATTCAG 1567

QY 472 AspLeuAlaSerGluLeuValHisTrpGluPheLeuHisGluAspArgMetLysLeu 491
DB 1568 GAGCTGGGAGCAGAGACTGTTACAGTTGGCTTTGACAGTGTCCGACCAATTTGACATC 1627

QY 492 AlaAlaPheLeuGluSerThrPheLeuLysTrp 502
DB 1628 GCCTGTTAATTGAAGAACTTTCAAGCAATTC 1660

RESULT 15
BC004756 2128 bp mRNA linear ROD 29-JUN-2004
LOCUS Mub muscular nuclear receptor binding protein, mRNA (cdna clone
DEFINITION MG:6961 IMAGE:3154089), complete cds.
ACCESSION BC004756
VERSION BC004756.1 GI:13435803
KEYWORDS MGC, muscle (house mouse)
SOURCE

ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 2128) Straubner, R.D., Collins, F.S., Wagner, L., Shemen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buelow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datschenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stadler, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Udell, T.B., Toshitsuki, S., Carminci, P., Prange, C., Kaha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McKwan, P.J., McKeown, R.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Viallon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bonfield, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, J., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	12477932
REFERENCE	2 (bases 1 to 2128)
AUTHORS	Straubner, R.
TITLE	Direct Submission
JOURNAL	Submitted (21-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgasbs-remail.nih.gov Tissue Procurement: Lochar Hennighausen Ph.D., Chu-Xia Deng Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amgdbcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisged, H., Kowit, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Navevati, A.N., Gibbs, R.A.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at: http://image.llnl.gov Series: IRAX Plate: 6 Row: a Column: 1 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22219433. Location/Qualifiers 1..2128 /organism="Mus musculus" /strain="129,C57BL/6T, FVB/N" /db_xref="taxon:10090" /clone="MGC:6961 IMAGE:3154089" /tissue_type="Mammary tumor. Brcal-1/EL; MMTV-Cre model. 10 months old, gross tissue." /clone_id="NCI CGAP_Mam3" /lab_host="DH10B" /note="vector: pCMV-SPORT6" 1..2128 /gene="Ntrp" /note="synonym: MGC6961" /db_xref="GeneID:192292" /db_xref="MGI:2183436" 75..1682 /gene="Ntrp"
gene	
CDS	

[illegible]

Db	822	GTCAACAAACGTGACACACAGCTGGACATCTTATCTCTTTGGCATGTGCACTGGAGATG	881
Qy	256	AlaValLeuGluIleGlnThrAsnGlyAspThrArg---ValThrGluGluAlaIleAla	274
Db	882	GCAGTGTCTGGAGATTCAAGGGCAATGGCAGATCTCATATGTGCCACAGAAAGCCATCAAC	941
Qy	275	ArgAlaArgHisSerLeuSerAspProAsnMetArgGluPheIleLeuCysValLeuAla	294
Db	942	AGTGCATATCCAGCTACTAGAGACTCATTTACAGAGGAGATTATCAAAAGTGCCTGCAG	10010
Qy	295	ArgAspProAlaArgArgProSerLHisSerLeuLeuPheHisArgValLeuPheGlu	314
Db	1002	TCTGAGCTGTCTGGAGACCAACACCCAGAACTTGTGTTCCACCCAGCACTGTTTAA	10610
Qy	315	ValHisSerLeuValLeuLeuAlaAlaHisCysPheIleGlnHisGlnThrLeuMetPro	334
Db	1062	GTGGCCCTCACTCAAGCTTCTTGCTCTCACTGTATGCTGGGGACCAACATATATCCCA	11210
Qy	335	GluAsnValValGluGluValThrValaMetAspLeuHisValaValLeuAlaGluLeu	354
Db	1122	GAGAAACGCTCTAGAGAGATGATCCCAAGAAACATGATACAGATGCTGTACTAGCTGAATT	11810
Qy	355	ProArg---ProArgArgProProLeuGlnThrArgGlyTyrSerGluValSerPheMetGlu	373
Db	1182	CCCCGAGGGCCGAGACGAGAACCAAGTTCAAGCTTGTACTCTCACTCAACGCCCTTAAGA	12410
Qy	374	LeuAspArgPheLeuGluAspValaArgAsnGlyIleTyrProLeuMetAsnPheAlaAla	393
Db	1242	TTAGACAAATTCCTTAGAGATGTCAGGAATGGAGATCACTCCCTGCACAGCCTTT-----	12990
Qy	394	ThrArgProLeuGlyLeuProArgValLeuAlaProProGluGluVal-----	410
Db	1296	-----GGGCTACTCTGGGCTCAGACGCCACAGACAGAGAGATGATCAATCACT	13430
Qy	411	-----GlnValValLeuThrProThrProGluProPheAspSerGluThrArgGly	427
Db	1344	GTTGTGCCCCCTCTGTCAAGCTTCAACTCTGTGACCAAGCTGAAAGTGAAGACAGAAAG	14030
Qy	428	ValIleGlnMetGlnCysAsnLeuGluArgSerGluAspArgValaArgTyrHisIleLeuThr	447
Db	1404	GTGGTCTGATGCACTGCACTGCACATCTGTGGAGAGGAGATGCACACACCATTTAA	14630
Qy	448	LeuLeuLeuValLeuGluAspArgGlyHisArgGlnLeuThrTyrAspLeuLeuProThr	467
Db	1464	CTTCTGCTGAAGCTGGAGGACAAATTGAACCGGACCTGACCTGACTGATATGCCAAAT	15230
Qy	468	AspSerAlaGlnAspLeuAlaSerGluLeuValHisTyrGlyPheLeuHisGluAspAsp	487
Db	1524	GAGAGCAATCCCGGACTTGGCAGCTGAGCTGGTGGACAGCTGAGGCTTCATTAGAGAGGCTGAT	15830
Qy	488	ArgMetLeuValLeuAlaPheLeuGluSerThrPheLeuLeuTyrTyrArgGlyTyrGln	506
Db	1584	CAGAGCCGCTGACTTCTCTGCTGGAGAGACGCTCAACAAATTTCAATTTCCACAGG	16400

THIS PAGE BLANK (USPTO)


```

Db      155 CGGAGCAGGTAAACCAAGGAAACATCGCAGGGCTTCAAGACCTTCTAGCCATGAGAC 214
Qy      61  ThrGluGluGlyValGluValValTTPaangluLeuHisPheGlyAspArgIysAlaPhe 80
Db      215 ACGGAGAGAGGGGTGAGAGGTGGTGTGGAACGAGCTCACTCGGAGACAGAAAGGCTTC 274
Qy      81  AlaAlaHISgluGluLeuIleGlnThrValPheGluGlnLeuValLeuValAspHisPro 100
Db      275 GCGGGGACAGAGAGAGATCCAGACCGGTTCGAGACCTGGTGTGGAACCAACCGG 334
Qy      101 AsnIleValIysLeuHisIleYsYrTTPLeuAspThrSerGluAlaCysAlaArgValIle 120
Db      335 AACATCGTGAAGTTCACAAATACGCTGAGATACCTTCGAGGCGCTCGCAGAGGTCATC 394
Qy      121 PheIleThrGluTyrValSerSerGlySerIleuValGlnPheLeuIlyIysThrIysIys 140
Db      395 TTCATCACAGATAGCTGTCTATCAGGACCTCAAGCAATTCCTCAAAAGACCAAGAG 454
Qy      141 AsnHisIysValMetAsnAlaArgAlaTTPlyAsArgTTPCysThrGlnIleLeuSerAla 160
Db      455 AACCAAGAGCCATGAACGCCCGGCTGAAAGCGGTGACACGCAATCCTGTCTGG 514
Qy      161 LeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAsnLeuThrSerAspThr 180
Db      515 CTCAGCTTCCTGCAGCGCTGCAGCCCCCAATCATCAAGGAACTTGACCAAGCAACAC 574
Qy      181 IlePheIleGlnHisAsnGlyLeuIleIysIleGlySerValTTPHisArgIlePheSer 200
Db      575 ATCTTCATTCAGCAACACGGCTCATCAAGATCGGCTCGGTGGCAACGAATCTTCTCC 634
Qy      201 AsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIleArgIleGlu 220
Db      635 AAT-----GCACTCCAGATGATCTCGAAGCCCATCCGCGCTGAG 676
Qy      221 ArgGluGluLeuArgAsnLeuHisPhePheProProGluTyrGlyGluValAlaAspGly 240
Db      677 CGAGAGGAACCTCGGAACCTGCATCTTCCCCCAAGATAGAGAGGTGGCCCATGGG 736
Qy      241 ThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGlnMetAlaValLeuGluIle 260
Db      737 ACCGCTGTGACATCTTCTTCTTGGAGTGTGCGCTGAGATGGCTGTACTGGAATTC 796
Qy      261 GlnThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArgHisSerLeu 280
Db      797 CAGACCAATGGGGAACCCGGGTCAAGAGAGGCCATTCCTGGCCGACGCACTCGCTG 856
Qy      281 SerAspProAspMetArgGluPheIleLeuCysValLeuAlaArgAspProAlaArgArg 300
Db      857 AGTGACCCCAACATCGGAGATTCTTCTGTGCTGTGGCCGGGAACTTCCCGACG 916
Qy      301 ProSerAlaHisSerLeuLeuPheHis-ArgValIleuPheGluValHisSerLeuIysLe 320
Db      917 CCTCTGTGTCACAGCTCTCTTCCACNCGGCTGCTTNGAGGTGCACTCGCTGAAGCT 976
Qy      320 uleu 321
Db      977 CCTG 980

```

RESULT 3

```

US-09-862-027-9
Sequence 9, Application US/09862027
Patent No. 6858418
GENERAL INFORMATION:
APPLICANT: Hodge, Martin R.
TITLE OF INVENTION: No. 6858418el Kinases and Uses Thereof
FILE REFERENCE: 35800/234862
CURRENT APPLICATION NUMBER: US/09/862,027
CURRENT FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 09/345,473
PRIOR FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 82
SOFTWARE: FASTSEQ for Windows Version 4.0

```

```

; SEQ ID NO 9
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ... (981)
; OTHER INFORMATION: n=A,T,C or G
US-09-862-027-9

Alignment Scores:
Pred. No.: 2,52e-169 Length: 981
Score: 1617.00 Matches: 312
Percent Similarity: 96.89% Conservative: 0
Best Local Similarity: 96.89% Mismatches: 3
Query Match: 60.56% Indels: 7
DB: 3 Gaps: 1

US-10-618-941-99 (1-507) x US-09-862-027-9 (1-981)
Qy      1  MetAlaIAProGluProAlaProArgAlaArgGluArgGluArgGluArg 20
Db      35  ATGGGCGCCCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 94
Qy      21  GluSerGluAspGluSerAspIleLeuGluGluSerProCysGlyArgTTPGlnIysArg 40
Db      95  GAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 154
Qy      41  ArgGluGlnValAsnGlnIysAsnMetProGlyLeuGlnSerThrPheLeuAlaMetAsp 60
Db      155  CGGAGCAGGTAAACCAAGGAAACATGCGAGGCTTCAGAGACCTTCTAGCCATGAGAC 214
Qy      61  ThrGluGluGlyValGluValValTTPaangluLeuHisPheGlyAspArgIysAlaPhe 80
Db      215  ACGGAGAGAGGGGTGAGAGGTGGTGTGGAACGAGCTCACTCGGAGACAGAAAGGCTTC 274
Qy      81  AlaAlaHISgluGluLeuIleGlnThrValPheGluGlnLeuValLeuValAspHisPro 100
Db      275  GCGGGGACAGAGAGAGATCCAGACCGGTTCGAGACCTGGTGTGGAACCAACCGG 334
Qy      101  AsnIleValIysLeuHisIleYsYrTTPLeuAspThrSerGluAlaCysAlaArgValIle 120
Db      335  AACATCGTGAAGTTCACAAATACGCTGAGATACCTTCGAGGCGCTCGCAGAGGTCATC 394
Qy      121  PheIleThrGluTyrValSerSerGlySerIleuValGlnPheLeuIlyIysThrIysIys 140
Db      395  TTCATCACAGATAGCTGTCTATCAGGACCTCAAGCAATTCCTCAAAAGACCAAGAG 454
Qy      141  AsnHisIysValMetAsnAlaArgAlaTTPlyAsArgTTPCysThrGlnIleLeuSerAla 160
Db      455  AACCAAGAGCCATGAACGCCCGGCTGAAAGCGCTGTGACCAACAGATCTCTGTGCG 514
Qy      161  LeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAsnLeuThrSerAspThr 180
Db      515  CTCAGCTTCCTGCAGCGCTGCAGCCCCCAATCATCAAGGAACTTGACCAAGCAACAC 574
Qy      181  IlePheIleGlnHisAsnGlyLeuIleIysIleGlySerValTTPHisArgIlePheSer 200
Db      575  ATCTTCATTCAGCAACACGGCTCATCAAGATCGGCTCGGTGGCAACGAATCTTCTCC 634
Qy      201  AsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIleArgIleGlu 220
Db      635  AAT-----GCACTTCAGATGATCTCGAAGCCCATCCGCGCTGAG 676
Qy      221  ArgGluGluLeuArgAsnLeuHisPhePheProProGluTyrGlyGluValAlaAspGly 240
Db      677  CGAGAGGAACCTCGGAACCTGCATCTTCCCCCAAGATAGAGAGGTGGCCCATGGG 736
Qy      241  ThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGlnMetAlaValLeuGluIle 260
Db      737  ACCGCTGTGACATCTTCTTCTTGGAGTGTGCGCTGAGATGGCTGTACTGGAATTC 796
Qy      261  GlnThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArgHisSerLeu 280

```

DB 797 CAGACATGGGAGACCCGGCTCACAGAGAGCCATTGCTCGGCCAGGACTCGCTG 856
QY 261 SerAapProAmeMetArgGluPheIleuCySyleuAlaAArgAapProAlaArgArg 300
DB 857 AGTGACCCCAACATCGGGAGTTTCATCTTCTGCTGCGCCGGAGACCTGCCGAGCG 916
QY 301 ProSerAlaHisSerIleuPheHis-ArgValIleuPheGluValHisSerIleuLysLe 320
DB 917 CCCTGTGTCACAGCCTCTCTTCCACACGCGTGTCTTTNGAGGTGCACTGCTGAAGCT 976
QY 320 uLeu 321
DB 977 CCTG 980
RESULT 4
US-08-933-750C-69
; Sequence 69, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2029 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITUT03
; CLONE: 864683
; US-08-933-750C-69
Alignment Scores:
Pred. No.: 1.13e-159 Length: 2029
Score: 1534.50 Matches: 305
Percent Similarity: 73.54% Conservative: 59
Best Local Similarity: 61.62% Mismatches: 102

Query Match: 57.47% Indels: 29
DB: 2 Gaps: 6
US-10-618-941-99 (1-507) x US-08-933-750C-69 (1-2029)
QY 17 GUAArgGUAePrgUSeRglUAaPglUSeRapiLleuGluGluSeRProCysEglYArg 36
DB 240 GAGAAAGAAAGAAAGTGAAGATGATGATGATTTTGGAAAGTCCCTGTGGGCGC 299
QY 37 TrrGlnYsAArgArgGluGlnValAsnGlnYAsnMetProGlyLeuInSerThrPhe 56
DB 300 TGGCAGAAAGAGCCAGAAAGAGTGAATCAACGAAATGATACAGGATATGACAGTGCATAC 359
QY 57 LeuAlaMetAePrrGluGluGlyValGluValValTrrPaenGluLeuHisPheGlyAap 76
DB 360 CTGGCCATGATGATACAGAGAAAGGTGTAGAGGTGTGTGAATGAGATACAGTTCCTGA 419
QY 77 ArgLYsAlAPheAlaHisGluGluLysIleGlnThrValIPheGluGlnLeuValLeu 96
DB 420 CGCAAGAACTCAAGCTGCAGAGAGAAAGAGTGTGCTGTGTGATTAATCTGATTCAA 479
QY 97 ValAspHisProAenIleValIysLeuHisLYsTYrTrrLeuAapThrSerGluAlaCys 116
DB 480 TTGAGGATCTTAATCATTTGTTAATTCACAAATATTGGGCTGACATTAAAGAGAACAG 539
QY 117 AlaArgValIlePheIleThrGluTYrValSerSerGlySerLeuLYsGlnPheLeuLYs 136
DB 540 GCCAGGTCATTTTATTCACAGAAATACATGCATCTGGAGTCTGAAAGCAATTTCTGAAG 599
QY 137 LYsThrLYsLYsAenHisLYsValAMetAsnAlaArgAlaTrrLYsAArgTrrCysThrGln 156
DB 600 AAGACCAAAAGAAACCAAGACATGATATAAAGCATGAAGCGTTGGTGCACACAA 659
QY 157 IleLeuSerAlaLeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAsnLeu 176
DB 660 ATCTCTCTGCTGCTTAAGTACTGCTGACCTGCTGACCCCATCATCATCGGAACCTG 719
QY 177 ThSerAapThrIlePheIleGlnHisAsnGlyLeuIleLYsIleGlySerValTrrHis 196
DB 720 ACCGTGACACCATTCATTCACAGACACGAGCTCATCAAGATTGGCTGTGTG----- 773
QY 197 ArgIlePheSerAenAlaLeuArgProProThrAlaLeuPrrAaPrrLeuAArgSerPro 216
DB 774 -----GCTCTGACACTATCAACATAT 797
QY 217 IleArgAlaGluArgGluGluLeuArgAsnLeuHisPhePheProGluTYrGlyGlu 236
DB 798 GTGAAGACTGTGCGAAGAGAGCAAGAAATCTACCTTTTGGACCAAGATAGAGAA 857
QY 237 ValAlaAap---GlyThrAlaValAapIlePheSerPheGlyMetCysAlaLeuGluMet 255
DB 858 GTCACTATGTCAGAAAGAGAGTGAATCATCTCTTTGGCATGTGTGCACTGAGATG 917
QY 256 AlaValIleuGluIleGlnThrAsnGlyAapThrArg---ValTrrGluGluAlaIleAla 274
DB 918 GCAGTGTGAGATTCAGGGCAATGAGAGCTTCATATGTGCCACAGAGGCCATCGC 977
QY 275 ArgAlaArgHisSerIleuSerAapProAmeMetArgGluPheIleLeuCYsCysLeuAla 294
DB 978 AGTGCATCCAGCTTCTTGAAGAGCCATTACAGAGGAGATTCACTTCAAAATGCTCGAG 1037
QY 295 ArgAapProAlaArgArgProSerAlaHisSerIleuPheHisArgValIleuPheGlu 314
DB 1038 TCTGAGCCTGTGTCAGAACCAACGACAGAACTTCTGTTCACACCGACATTTGTTGA 1097
QY 315 ValHisSerIleuYsLeuAlaHisGlyPheIleGlnHisGlnTYrLeuMetPro 334
DB 1098 GTGCCCTGTGCTCAAACTCTTGCGCCACCTGCACTTGTGGACCAACACATATATCCCA 1157
QY 335 GUAaenAlaValGluGluLYsThrLYsAlAMetAePrrLeuHisAlaValIleuAlaGluLeu 354
DB 1158 GAGAAACGCTTGAAGAGATCACCAAAACATGATGATGTCGCTGCTGCTGAATC 1217


```
Oy 275 Arg1aAArgHisSerLeuSerAspProAsnMetArgGluPhe1LeuCyCysLeuAla 294
Db 978 AGTGCATCCAGCTTCTTGAAAGACCATTCACAGAGGAGTTCAATCAAAAGGCTGGAG 1037
Oy 295 ArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGlu 314
Db 1038 TCTGAGCCTGGCTCGAGACCAACAGCCAGAGAACTTCTGTTCCACCCAGCATTTGTTGAA 1097
Oy 315 ValHisSerLeuValLeuLeuAlaHisCysPheHisLeuHisGlnTyrLeuMetPro 334
Db 1098 GTGCCCTGCTCAAACTCTTGGCGCCACATGTCATCTGGGACACCAACATGATCTCCA 1157
Oy 335 GluAsnValValGluGluValThrValAlaMetAspLeuHisAlaValLeuAlaGluLeu 354
Db 1158 GAGAACGCTCTAGAGGAGATCACCAAAAACATGATCTAGTCCGTACTGAGTGAATTC 1217
Oy 355 ProArg---ProArgArgProProLeuGlnTyrArgTyrSerGluValSerPheMetGlu 373
Db 1218 CCTGAGGACCAAGAGAGAACCACTTCAGACTTGTACTCTCACTCAACAGCTGTGAA 1277
Oy 374 LeuAspLysPheLeuGluAspValArgAenglyIleTyrProLeuMetAsnPheAlaAla 393
Db 1278 TTAGATAAATTCCTTGAAAGATGTCAGAAATGGGATCTATCTCTGACAGCCTTT----- 1331
Oy 394 ThrArgProLeuGluValProArgValLeuAlaProProGluGluVal----- 410
Db 1332 -----GGGCTCTCTCGGCCCCACGACCCACGACGAGAGAGAGTACATCACCT 1379
Oy 411 -----GlnLysAlaLysThrProThrProGluProPheAspSerGluThrArgLys 427
Db 1380 GTCCGAGCCCTCTGTCAAGATCTCCGACACTGAAACAGCTGAGTGAAGACTGGCAG 1439
Oy 428 ValIleGlnMetGlnCysAsnLeuGluArgSerGluAspValAlaArgTyrHisLeuThr 447
Db 1440 GTGGTCTGATGTCAGTGCACAACTGATGTCGGTGGAGGAGGAGTCAAAACACACTGGACA 1499
Oy 448 LeuLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThrTyrAspLeuLeuProThr 467
Db 1500 CTCTGCTGAAGTTGGAGGACAACTGAACCGGACCTGAGCTGTGACTGATGCAAT 1559
Oy 468 AspSerAlaGlnAspLeuAlaSerGluLeuValHisTyrGlyPheLeuHisGluAspAsp 487
Db 1560 GAGAAATATCCCGAGAGTTGCGGCTGAGCTGGTGCAGCTGGGCTTCATATGAGAGCTGAC 1619
Oy 488 ArgMetLysLeuAlaAlaPheLeuGluSerThrPheLeuTyr 502
Db 1620 CAGAGCCGTTGACTTCTCTGCTAGAGAGACCTTGAAACAAGTTC 1664

RESULT 6
US-09-949-016-1428
; Sequence 1428, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241, 755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237, 768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231, 498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1428
; LENGTH: 2163
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1428
```

```
Alignment Scores:
Pred. No.: 1.26e-159 Length: 2163
Score: 1534.50 Matches: 305
Percent Similarity: 73.54% Conservative: 59
Best Local Similarity: 61.62% Mismatches: 102
Query Match: 57.47% Indels: 29
DB: 3 Gaps: 6

US-10-618-941-99 (1-507) x US-09-949-016-1428 (1-2163)
Oy 17 GluArgGluAspGluSerGluAspGluSerAspIleLeuGluGluSerProCysGlyArg 36
Db 242 GAGGAAAGAAAGAAAGTGAAGATGATCTCAAGATTTTGGAAAGTCCCTCTGGCGGC 301
Oy 37 TrrGlnLysArgArgGluGluValAsnGlnGlyAsnMetProGlyLeuGlnSerThrPhe 56
Db 302 TGGCAGAAAGAGGCGAGAAAGAGTGAATCAACGGAAATGTACAGAGGTATTGACATGCATAC 361
Oy 57 LeuAlaMetAspThrArgGluGluGluValAlaGluValAlaTrrAsnGluLeuHisPheGlyAsp 76
Db 362 CTGGCCATGTGATACAGAGAAAGGTGTAGAGTTGTGTGAATGTAGATCAAGTCTCTGAA 421
Oy 77 ArgLysAlaPheAlaAlaHisGlnGluLysIleGlnThrValPheGluGlnLeuValLeu 96
Db 422 CGCAAGAACTACAAAGCTCGACAGAGAAAGAGTTGCTGTGTTGATATCTGATTCAA 481
Oy 97 ValAspHisProAsnIleValLysLeuHisLysTyrTrrLeuAspThrSerGluAlaCys 116
Db 482 TTGAGCATCTTAACATGTTGAATTTACAAATATTTGGGTGATTAAGAGAAACAG 541
Oy 117 AlaArgValIlePheIleThrGluTyrValSerSerGlySerLeuLysGlnPheLeuLys 136
Db 542 GCCAGGTCATTTTATACAGAAATACATGTCATCTGGAGTCTGAAGCAATTTCTGAAG 601
Oy 137 LysThrLysLysAsnHisLysValAlaMetAsnAlaArgAlaTrrLysArgTrrCysThrGln 156
Db 602 AAGACCAAAAGAAACCAAGACATGAATGAAGGACATGGAAGCTGTGTCACAA 661
Oy 157 IleLeuSerAlaLeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAsnLeu 176
Db 662 ATCTCTCTGCCCTTAACCTTCACTGCACTCCGTTGACCCCTCATTCATCAGGAACTCG 721
Oy 177 ThrSerAspThrIlePheIleGlnHisAsnGlyLeuLeuLysIleGlySerValTrrHis 196
Db 722 ACCTGTACACCACTTTCATCCAGACAGAGTCAATCAAGATGCTGTG----- 775
Oy 197 ArgIlePheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerPro 216
Db 776 -----GCTCTGACATATCAACATCAT 799
Oy 217 IleArgAlaGluArgGluGluLeuArgAsnLeuHisPhePheProProGluTyrGlyGlu 236
Db 800 GTGAAAGACTTGTGAGAGAGAGAGAAATCTTACACTTCTTGGACACAGAGTATGAGAA 859
Oy 237 ValAlaAsp---GlyTrrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMet 255
Db 860 GTCACTATGTGACAAAGAGAGTGAATCTACCTCTTGGCAGATGTGCACTGAGATGG 919
Oy 256 AlaValLeuGluLysIleGlnTrrAsnGlyAspThrArg---ValTrrGluGluAlaIleAla 274
Db 920 GCACTGCTGAGATTCAGGCGCAATGAGAGTCCATATGTGCCACAGAAAGCCATCAGC 979
Oy 275 ArgAlaArgHisSerLeuSerAspProAsnMetArgGluPheIleLeuCyCysLeuAla 294
Db 980 AGTGCATCCAGCTTCTTAAGAAAGCCATTCACAGAGGAGTTCATTAAGAGTGCCTGAG 1039
Oy 295 ArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGlu 314
Db 1040 TCTGAGCCTGCTCGAGACCAACAGCCAGAGAACTTCTGTTCCACCCAGCATTTGTTGAA 1099
Oy 315 ValHisSerLeuValLeuLeuAlaHisCysPheHisLeuHisGlnTyrLeuMetPro 334
```

Db 1100 GTGCCCCGCTGAACTCCTCGGCGCCACCTGATTTGGGACACCAACATGATCCCA 1159
Qy 335 GUAAGNValAGUGUlyrThryrValmetAapLeuH1AlaValleuAgluleu 354
Db 1160 GAGAACGCTTAGAGAGATACCAAAAATGATATCTAGTGGCTAGCTGGAATC 1219
Qy 355 ProArg---ProArgArgProProleuGlnTyrArgTyrSerGluValSerPheMetGlu 373
Db 1220 CCGCAGAGCCAGAAAGAACCAAGTTCAGACTTTGTACTTCTCAGTCCACAGCTCGGAA 1279
Qy 374 LeuAapLyPheLeuGluAapValArgaenglyllyrProleuMetAapPheAla 393
Db 1280 TTAGATAAATTCCTGAAATGTCAGAAATGGATCTATCTCTGACACCTTT----- 1333
Qy 394 ThrArgProLeuGluValProArgValLeuAlaProProProGluGluVal----- 410
Db 1334 -----GGGCTGCTCGGCGCCCGCCAGCCACAGAGAGAGAGGTGACATCCT 1381
Qy 411 -----GlnLyAlaLyThrProThrProGluProPheAapSerGluThraArglys 427
Db 1382 GTGCGCCCCCTGTGTCAAGATCCGACACCTGAACAGCTGAGGTGAGACTCGCAAG 1441
Qy 428 ValIleGlnMetGlnCyAaenLeuGluArgSerGluAapLyAlaArgTrrpHisLeuThr 447
Db 1442 GTGGTGTGATCGAGTCAACATGATGTCGTGAGAGAGATCAACACCACTGACA 1501
Qy 448 LeuLeuLeuValLeuGluAapArgLeuH1AargGlnLeuThrTyrAapLeuProThr 467
Db 1502 CTTCGCTGAGAGTGAAGCAACTGAACCGGACCTGAGCTGTGACTGATGCCAAAT 1561
Qy 468 AapSerAlaGlnAapLeuAaSerGluLeuValH1AryrGlyPheLeuH1AgluAap 487
Db 1562 GAGATATCCCGGAGTGGCGGCTGAGCTGCTGACGTGGCTTATTGTAAGCTGAC 1621
Qy 488 ArgMetLyLeuAlaAaPheLeuGluSerThrPheLeuLyArg 502
Db 1622 CAGACCGGTTGACTTCTGCTAGAAAGACCTTGAAACAGTTTC 1666
RESULT 7
US-09-023-942A-9
; Sequence 9, Application US/09023942A
; Patent No. 6479274
; GENERAL INFORMATION:
; APPLICANT: (US only) ANTALIS Toni Marie and HOOPER John David
; TITLE OF INVENTION: NOVEL MOLECULES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,942A
; FILING DATE: 13-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: POS101/97
; FILING DATE: 13-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP0422/97
; FILING DATE: 18-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: International PCT Application
; FILING DATE: 13-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGIGLIO, FRANK S

; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 11168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742 4343
; TELEFAX: (516) 742 4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2241 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 166..1773
; US-09-023-942A-9
Alignment Scores:
Pred. No.: 1,01e-156 Length: 2241
Score: 1508.50 Matches: 301
Percent Similarity: 72.93% Conservative: 60
Best Local Similarity: 60.81% Mismatches: 105
Query Match: 56.50% Indels: 29
DB: Gaps: 6
US-10-618-941-99 (1-507) x US-09-023-942A-9 (1-2241)
Qy 17 GUAArgGluAapGluSerGluAapGluSerAapIleLeuGluGluSerProCyArg 36
Db 295 GAGGAAGAAGAAAGAAAGTGAATGATGCTGAGATTTGGAAAGATCGCCCTGTGGCGCC 354
Qy 37 TrrpGlnLyAaArgGluGlnValAengGlnGlyAaMetProGlyLeuGlnSerThrPhe 56
Db 355 TGGCAGAAAGCGCGAAGAGGTGATCAACCGAAATGTACCGATTAACAGTGCATC 414
Qy 57 LeuAlaMetAapThrGluGluGluValAlaGluValAlaTrrpAaGluLeuH1AapPheGlyAap 76
Db 415 CTGGCCATGATACAGAGAGAGTGAAGGTGATGAGTGTGAGATGAAGTACATTCCTGCA 474
Qy 77 ArgLyAlaAaPheAlaAaH1AgluGluLyAaIleGlnThrValPheGluGlnLeuValleu 96
Db 475 CGCAAGACATCAACACTGCGAGAGAGAAAGCTTGGCTGTGATTAATTGATTCA 534
Qy 97 ValAaPheAaProAaH1VallyLeuH1AlyrTrrpLeuAapThrSerGluAaCyS 116
Db 535 TTGGAGCATCTTAACATTTGTTAACTTTCACAAATTTGGCTGACATTAAAGAAACAAG 594
Qy 117 AlaArgValIlePheAlleThrGluLyAaSerArgLySerLeuLyGlnPheLeuLyS 136
Db 595 GCCAGGGTCAATTTTATCAAGATACATGATCTGGAGTCTGAAGCAATTTCTGAAG 654
Qy 137 LySThrLyAaSerH1LyAaLeuAaMetAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 156
Db 655 AAGAACCAAAAGAACCAACAGCATGAATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 714
Qy 157 IleLeuSerAlaLeuSerPheLeuH1AaCySerProPheIleIleHisGlyAaenLeu 176
Db 715 ATCCCTCTGCGCCCTTAAGCTACCTGCACTCTGTGACCCCATCATCATGGAACCTG 774
Qy 177 ThrSerAapThrIlePheAlleGlnH1AaengLyLeuIleLyAaIleGlySerAlaTrrpHis 196
Db 775 ACCTGTGACACCATCTTATCCAGCAACGACATCAAGATTTGGCTGTG----- 828
Qy 197 ArgIlePheSerAaAlaLeuAaArgProThrAlaLeuProAaPheLeuAaArgSerPro 216
Db 829 -----GCTCTGACACATCAACATCAT 852
Qy 217 IleArgAlaGluArgGluGluLeuAaenLeuH1AaPhePheProProGluTrrpGlyGlu 236
Db 853 GTGAAGACTTGGAGAGAGAGCAAGAAATCTTACCTTTGACACAGATGAAGAA 912
Qy 237 ValAlaAap---GlyThrAlaValAaPheIlePheSerPheGlyMetCyAlaLeuGluMet 255

```
Db 913 GTCACTAATGTGACAAAGCAGTGCATCTACTCTTGGATGTGTGCATGGCGATG 972
Qy 256 A1AvalLeuGlu11eGlnThrAsnGluAspThrArg---ValThrGluGluAla11eAla 274
Db 973 GCAGGCTGTGAGATTCCAGGCGCAATGAGATGCTCATATGTGTCCACAGAAAGCCATCAGC 1032
Qy 275 Arg1AaArgH1AspLeuSerAspProAsnMetCArgu11eLeuGluCysGluAla 294
Db 1033 AGTGCATCCAGCTTCTAGAAAGACCTTACAGAGGAGTTCATCAAAAGTGCCTGCAG 1092
Qy 295 ArgAspProAlaArgArgProSerAlaH1AspLeuLeuPheH1AspValLeuPheGlu 314
Db 1093 TCTGAGCCTGTGCGAGACCAACAGCAGAGAACTTCTGTTCCACCAGCATGTTTGA 1152
Qy 315 ValH1AspLeuLeuLeuAlaH1AspPheH1eGlnH1eGlnThrLeuMetPro 334
Db 1153 GTGCCCTGCTCAAACTCTTCCGCGCCCATGTCATTTGGGACACCAACATGATCCCA 1212
Qy 335 GluAsnValValGluGluLeuThrValMetAspLeuH1AspValLeuAlaGluLeu 354
Db 1213 GAGAAAGCTCTAGAGAGATCCCAAAACATGATACTAGTGCCTGATCGCTGAATTC 1272
Qy 355 ProArg---ProArgArgProProLeuGlnThrArgTyrSerGluValSerPheMetGlu 373
Db 1273 CCTGAGAGACCAAGAGAAACAGTTCAGACTTGTGACTCTGACGTACAGCTCTGAA 1332
Qy 374 LeuAspLysPheLeuGluAspValArgAsnGly11eTyrProLeuMetAsnPheAla11e 393
Db 1333 TTAGATAAATTCCTTGAAGATGTCAAGATGGATGATCTATCTCTGACAGCCCTT 1386
Qy 394 ThrArgProLeuGlyLeuProArgValLeuAlaProProProGluGluVal 410
Db 1387 -----GGGCTGCTCGGCCCAAGCAGCAGCAGAGAGAGATGACATCACCCT 1434
Qy 411 -----GlnLysAlaLeuThrProThrProGluProPheAspSerGluThrArgLys 427
Db 1435 GTCCGAGCCCTCTGTCAAGACTCCGACACTGAAACAGAGCTGAGAGACTGCGAAG 1494
Qy 428 Val11eGlnMetGlnCysAsnLeuGluArgSerGluAspLysValAspTyrH1eLeuThr 447
Db 1495 GTGGCTGAGTCAAGTGCACATTTGATGTCGGTGGAGGAGTCAAAACACACCTGCACA 1554
Qy 448 LeuLeuLeuValLeuGluAspArgLeuH1AspGlnLeuThrTyrAspLeuLeuProThr 467
Db 1555 CTTGCTGAAAGTGGAGACAACTGAACCGGCACTGAGCTGTGACTGATGCAAAAT 1614
Qy 468 AspSerAlaGlnAspLeuAlaSerGluLeuValH1eTyrGlyPheLeuH1eGluAspAsp 487
Db 1615 GAGAAATATCCCGAGATTGCGCGCTGAGCTGTGAGCTGGGCTTCATTATGAGGCTGAC 1674
Qy 488 ArgMetLysLeuAlaAlaPheLeuGluSerThrPheLeuLysTyr 502
Db 1675 CAGAGCCGCTGACTTCTGTGTAAGAGACCTTGAACAAGTTC 1719

RESULT 8
US-10-104-047-1939
; Sequence 1939, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104.047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1939
; LENGTH: 2546
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
US-10-104-047-1939
Alignment Scores:
Pred. No.: 9.66-44 Length: 2546
Score: 490.50 Matches: 115
Percent Similarity: 55.24% Conservative: 1
Best Local Similarity: 54.76% Mismatches: 1
Query Match: 18.37% Indels: 94
Dbs: 3 Gaps: 3

US-10-618-941-99 (1-507) x US-10-104-047-1939 (1-2546)
Qy 391 PheAla1aThrArgProLeuGluLeuProArgValLeuAlaProProProGluGluVal 410
Db 2 TTGCAAGCTTCAACCCCTGGAGCTGCCGTGTGCTGGCCCAACCCCGAGAGGTC 61
Qy 411 GlnLysAlaLysThrProThrProGluProPheAspSerGluThrArgLysVal11eGln 430
Db 62 CAAAAGGCAAGACCCGAGCGCCAGAGCCCTTGACTGTGAGACCAAGAAAGTCAATCAG 121
Qy 431 MetGlnCysAsnLeuGluArgSerGluAspLysAlaArgTyrH1s----- 445
Db 122 ATGCAATGCAACTGTGAGAGAAAGCAGAGACAGCGCGCTGAGATGAGCGGCGCTGG 181
Qy 445 ----- 445
Db 182 GAGGCGGCGCTGCGCTGGGAGAGCGGAGCGGAGCCCTGAACCCCGAGCGAC 241
Qy 446 -----LeuThrLeuLeuLeuValLeuGluAspArgLeuH1sArg 458
Db 242 CAGCGCCCGCGCTCCCAAGCTCATCTGCTTGTGGTGTGGAAGACGGCTGACCGG 301
Qy 459 GlnLeuThrTyrAspLeuPro----- 466
Db 302 CAGCTGACTTACGACCTGCTCCAAATGAGCTGGGACAGGACTTGGCGCGCGCGTGG 361
Qy 467 -----ThrAs 468
Db 362 ACCCGAGCGGCGGAGGAGCGGAGCGGCTCATGCTCGCTTACGCCATTGCTGACGGA 421
Qy 468 pSerAlaGlnAspLeuAlaSerGluLeuValH1eTyrGlyPheLeuH1eGln----- 485
Db 422 CAGCGCCAGGACCTGCGCTCGAGCTGTGCACTATGGCTCTTCACGA-GGTGGCGCT 480
Qy 485 ----- 485
Db 481 GGGCGGTGCGCGCGCTGCGGAGGAGGCGCAAGCGGCGAGCGCGCCCTCGTCC 540
Qy 486 -----AspAspArgMetLysLeuAla1aPheLeuGluSerTh 498
Db 541 CCATGCTCCCTCTTCCGCAAGACGAGATGAGCTGCGCGCTTCTGAGAGACAC 600
Qy 498 rPheLeuLysTyrArgLysThrGlnAla 507
Db 601 CTTCTCAATGATCCGTGGAGCCAGGCC 628

RESULT 9
US-09-949-016-13170
; Sequence 13170, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949.016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
```



```

GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Wang, Tonglong
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS
TITLE OF INVENTION: OF LUNG CANCER
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,984
FILING DATE: 18-MAR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.456
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-282-6031
TELEX:
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-040-984-23

Alignment Scores:
Pred. No.: 8,32e-33 Length: 669
Score: 383.50 Matches: 83
Percent Similarity: 66.24% Conservative: 21
Best Local Similarity: 52.87% Mismatches: 40
Query Match: 14.36% Indels: 13
Gaps: 3

US-10-618-941-99 (1-507) x US-09-040-984-23 (1-669)
Qy 349 AlAValleuAlaGlueuProAry---ProAryAryProProleuGIntPrAryTyrSer 367
Db 7 GCCGACTGGCTGAATCCCTGCAGACCAAGAGAACCAAGTTCAAGCTTTGACTCT 66
Qy 368 GluValSerPheMetGluLeuApylyPheLeuGluApyValAryAnglyIleTyrPro 387
Db 67 CAGTCACCAAGCTCTGAATTAGATAATTCCTTGAAGATGTCAGAAATGGATCTATCT 126
Qy 388 LeuMetAenPheAlaAlaThrAryProleuGlyLeuProAryValLeuAlaProProPro 407
Db 127 CTGACAGCTTT-----GGGCTGCTCGGCCCGCCAGCAGCAGCAGAG 168
Qy 408 GluGluVal-----GlnlyAlaYThrProThProThProGluProPhe 421
Db 169 GAGGAGGTGACATCACTGTCTGCCCCCTCTGCAAGACTCCAGACCTGAACCACT 228
Qy 422 AspSerGluThrAryGlyValIleGlnMetGlnCyAsnLeuGluArySerGluApylys 441
Db 229 GAGGTGAGAGACTCCGAAAGTGTCTGATGCACTGCAATTTGATCGGTGAGAGAGGA 288
Qy 442 AlaArgTrpHisLeuThrLeuLeuValLeuGluApyAryGluHisAryGlnLeuThr 461
Db 289 GTCAAAACCACTGACTTCTGCTGAAGTTGAGAGCAAACTGAACCGGACCTGAGC 348
Qy 462 TyrAspLeuLeuProThArySerAlaGlnApyLeuAlaSerGluLeuValHisTyrGly 481
```

```

Db 349 TGTGACCTGATGCCAAAGAGATATCCCGAGTTGGCGGCTGAGCTGTGACCTGGC 408
Qy 482 PheLeuHisGluApyApyAryMetIlyLeuAlaAlaPheLeuGlySerThr 498
Db 409 TTCTATTAGTGAAGCTGACCAAGCGGTTGACTTCTCTGTAGAAGAGACT 459

RESULT 11
US-09-123-912-23
Sequence 23, Application US/09123912A
Patent No. 6312695
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Wang, Tonglong
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
FILE REFERENCE: 210121.455C1
CURRENT APPLICATION NUMBER: US/09/123,912A
CURRENT FILING DATE: 1998-07-27
PRIOR APPLICATION NUMBER: 09/040,802
PRIOR FILING DATE: 1998-03-18
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 23
LENGTH: 669
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (642)
OTHER INFORMATION: where n is a, c, g or t
NAME/KEY: modified_base
LOCATION: (661)
OTHER INFORMATION: where n is a, c, g or t
US-09-123-912-23

Alignment Scores:
Pred. No.: 8,32e-33 Length: 669
Score: 383.50 Matches: 83
Percent Similarity: 66.24% Conservative: 21
Best Local Similarity: 52.87% Mismatches: 40
Query Match: 14.36% Indels: 13
Gaps: 3

US-10-618-941-99 (1-507) x US-09-123-912-23 (1-669)
Qy 349 AlAValleuAlaGlueuProAry---ProAryAryProProleuGIntPrAryTyrSer 367
Db 7 GCCGACTGGCTGAATCCCTGCAGACCAAGAGAACCAAGTTCAAGCTTTGACTCT 66
Qy 368 GluValSerPheMetGluLeuApylyPheLeuGluApyValAryAnglyIleTyrPro 387
Db 67 CAGTCACCAAGCTCTGAATTAGATAATTCCTTGAAGATGTCAGAAATGGATCTATCT 126
Qy 388 LeuMetAenPheAlaAlaThrAryProleuGlyLeuProAryValLeuAlaProProPro 407
Db 127 CTGACAGCTTT-----GGGCTGCTCGGCCCGCCAGCAGCAGCAGAG 168
Qy 408 GluGluVal-----GlnlyAlaYThrProThProThProGluProPhe 421
Db 169 GAGGAGGTGACATCACTGTCTGCCCCCTCTGCAAGACTCCAGACCTGAACCACT 228
Qy 422 AspSerGluThrAryGlyValIleGlnMetGlnCyAsnLeuGluArySerGluApylys 441
Db 229 GAGGTGAGAGACTCCGAAAGTGTCTGATGCACTGCAATTTGATCGGTGAGAGAGGA 288
Qy 442 AlaArgTrpHisLeuThrLeuLeuValLeuGluApyAryGluHisAryGlnLeuThr 461
Db 289 GTCAAAACCACTGACTTCTGCTGAAGTTGAGAGCAAACTGAACCGGACCTGAGC 348
Qy 462 TyrAspLeuLeuProThArySerAlaGlnApyLeuAlaSerGluLeuValHisTyrGly 481
Db 349 TGTGACCTGATGCCAAAGAGATATCCCGAGTTGGCGGCTGAGCTGTGACCTGGC 408
```

```

QY      482 PheLeuHLeGluAspArgMetLysLeuAlaA1PheLeuGluSerThr 498
DB      409 TTCATTAGTGAAGCTGACGACGAGCCGGTTCCTCTGCTGAAGAGACT 459

RESULT 12
US-09-643-597-23
/ Sequence 23, Application US/09643597
/ Patent No. 6426072
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongrong
/ APPLICANT: Fan, Liqun
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Hosken, Nancy
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Li, Samuel X.
/ APPLICANT: Wang, Aljun
/ APPLICANT: Skeiky, Yaelir A.W.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: McNeill, Patricia D.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.455C11
/ CURRENT APPLICATION NUMBER: US/09/643,597
/ CURRENT FILING DATE: 2000-08-21
/ NUMBER OF SEQ ID NOS: 369
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 23
/ LENGTH: 669
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(669)
/ OTHER INFORMATION: n = A,T,C or G
US-09-643-597-23

Alignment Scores:
Pred. No.:      8,32e-33      Length:      669
Score:          383.50      Matches:      83
Percent Similarity: 66.24%      Conservative: 21
Best Local Similarity: 52.87%      Mismatches: 40
Query Match:    14.36%      Indels:      13
DB:             3           Gaps:      3

US-10-618-941-99 (1-507) x US-09-643-597-23 (1-669)
QY      349 AlaValLeuAlaGluLeuProArg--ProArgArgProProLeuGluIntTPrArgTyrSer 367
DB      7 GCCGTACTGGCTGAATCCCTGACAGACAGAGAGAAACAGTTCAGCTTTGTACTCT 66
QY      368 GluValSerPheMetGluLeuAspLysPheLeuGluAspValArgAsnGlyIleTyrPro 387
DB      67 CAGTCACCAAGCTCTGGAATTGATTAATTCTTGAAGATGTCAGGATGGATCTATCT 126
QY      388 LeuMetAsnPheAlaAlaThrArgProLeuGlyLeuProArgValLeuAlaProProPro 407
DB      127 CTGACAGCTTT-----GGGCTGCTCGGCCCCAGCAGCAGCAGCAG 168
QY      408 GluGluVal-----GlnYsaIaLysThrProThrProGluProPhe 421
DB      169 GAGGAGGTGACATCACTGTCTGCTGCCCCCTGTCAACAGCTCCGACACTTAACCAAGCT 228
QY      422 AspSerGluThrArgLysValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLys 441
DB      229 GAGGTGAGACTCGCAAGGTGGTGATGATGCAATTGAGTGGTGAAGAGGGA 288
QY      442 AlaArgTrpHisLeuThrLeuLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThr 461
DB      289 GTCAAAACACCACTGACACTTCTGCGAAGTTGGAGCAAACTGAACCGGCACTGAGC 348
QY      462 TyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaAspSerGluLeuValHisTyrGly 481

```

```

DB      349 TGTGACTGATGCCAAATGAGATATCCCGAGTTGGCGGTGAGCTGGTGAAGTGGC 408
QY      482 PheLeuHLeGluAspArgMetLysLeuAlaA1PheLeuGluSerThr 498
DB      409 TTCATTAGTGAAGCTGACGACGAGCCGGTTCCTCTGCTGAAGAGACT 459

RESULT 13
US-09-480-884A-23
/ Sequence 23, Application US/09480884A
/ Patent No. 6482597
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongrong
/ APPLICANT: Fan, Liqun
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Hosken, Nancy A.
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Skeiky, Yaelir A.W.
/ APPLICANT: Henderson, Robert A.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.455C6
/ CURRENT APPLICATION NUMBER: US/09/480,884A
/ CURRENT FILING DATE: 2001-08-27
/ NUMBER OF SEQ ID NOS: 330
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 23
/ LENGTH: 669
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(669)
/ OTHER INFORMATION: n = A,T,C or G
US-09-480-884A-23

Alignment Scores:
Pred. No.:      8,32e-33      Length:      669
Score:          383.50      Matches:      83
Percent Similarity: 66.24%      Conservative: 21
Best Local Similarity: 52.87%      Mismatches: 40
Query Match:    14.36%      Indels:      13
DB:             3           Gaps:      3

US-10-618-941-99 (1-507) x US-09-480-884A-23 (1-669)
QY      349 AlaValLeuAlaGluLeuProArg--ProArgArgProProLeuGluIntTPrArgTyrSer 367
DB      7 GCCGTACTGGCTGAATCCCTGACAGACAGAGAGAAACAGTTCAGACTTTGTACTCT 66
QY      368 GluValSerPheMetGluLeuAspLysPheLeuGluAspValArgAsnGlyIleTyrPro 387
DB      67 CAGTCACCAAGCTCTGGAATTGATTAATTCTTGAAGATGTCAGGATGGATCTATCT 126
QY      388 LeuMetAsnPheAlaAlaThrArgProLeuGlyLeuProArgValLeuAlaProProPro 407
DB      127 CTGACAGCTTT-----GGGCTGCTCGGCCCCAGCAGCAGCAGCAG 168
QY      408 GluGluVal-----GlnYsaIaLysThrProThrProGluProPhe 421
DB      169 GAGGAGGTGACATCACTGTCTGCTGCCCCCTGTCAACAGCTCCGACACTTAACCAAGCT 228
QY      422 AspSerGluThrArgLysValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLys 441
DB      229 GAGGTGAGACTCGCAAGGTGGTGATGATGCAATTGAGTGGTGAAGAGGGA 288
QY      442 AlaArgTrpHisLeuThrLeuLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThr 461
DB      289 GTCAAAACACCACTGACACTTCTGCTGAAGTTGGAGCAAACTGAACCGGCACTGAGC 348
QY      462 TyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaAspSerGluLeuValHisTyrGly 481
DB      349 TGTGACTGATGCCAAATGAGATATCCCGAGTTGGCGGTGAGCTGGTGAAGTGGC 408
QY      482 PheLeuHLeGluAspArgMetLysLeuAlaA1PheLeuGluSerThr 498

```

```
Db      409 TTCATTAGTGAAGCTGACCAAGACCGGTTGACTTCTGTGTAAGAAGACT 459
RESULT 14
US-09-542-615A-23
; Sequence 23, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hoeken, Nancy A.
; APPLICANT: Panger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(669)
; OTHER INFORMATION: n = A,T,C or G
US-09-542-615A-23

Alignment Scores:
Pred. No.:      8,32e-33      Length:      669
Score:          383.50       Matches:      83
Percent Similarity: 66.24%   Conservative: 21
Best Local Similarity: 52.87% Mismatches: 40
Query Match:    14.36%      Indels:       13
DB:             3           Gaps:           3

US-10-618-941-99 (1-507) x US-09-542-615A-23 (1-669)
Qy      349 AAlaValIleuAlaGluLeuProArg--ProArgArgProProLeuGlnTrpArgTyrSer 367
Db      7   GCCGTACTGCTGAATATCCCTGCAGACACGAAAGAACCAAGTTCAACATTGTACTCT 66
Qy      368 GluValSerPheMetGluLeuAryPheLeuGluAryValAryGanGlyTleTyrPro 387
Db      67 CAGTCACCAAGCTCTGAATTAATTAATTCCTGAAGATGTCAGGAATGGGATCTATCT 126
Qy      388 LeuMetAnPheAlaAlaThrArgProLeuGlyLeuProArgValIleuAlaProProPro 407
Db      127 CTGACAGCCCTT-----GGGCTGGCTCGGCCCCCAGCCAGCCACAGCAG 168
Qy      408 GluGluVal-----GlnuYsaIalYsThrProThrProGluProPhe 421
Db      169 GAGAGAGGACATCACTGTCTGTCGCCCTCTGTCAAGACTCCGACACCTGAACCACT 228
Qy      422 AspSerGluThrArgLysValIleGlnMetGlnCyaaLeuGluAryGSerGluAryLys 441
Db      229 GAGGTGAGACTCGCAAGGTGGTCTGATGCACTGAACATTGAAGTCGATGGAAGAGGA 288
Qy      442 AlaArgTrpHisLeuThrLeuLeuValLeuGluAryAryGLeuHisArgGlnLeuThr 461
Db      289 GTCAAAACCAACCTGACATTCTGTGAAGTGAAGGACAACTGAACCGGACCTGAGC 348
Qy      462 TyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuValHisTyrGly 481
Db      349 TGTGACTGATGCCAAATGAAGATATCCCGAGTTGGCGGCTGAGCTGTGACGTGGGC 408
Qy      482 PheLeuHisGluAryAspAryGMetLysLeuAlaIaPheLeuGlnSerThr 498
Db      409 TTCAATTAGTAGGCTGACCAAGACCGGTTGACTTCTGTGCTAAGAAAGACT 459
,RESULT 15
```

```
US-09-606-421B-23
; Sequence 23, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hoeken, Nancy
; APPLICANT: Panger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasar A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(669)
; OTHER INFORMATION: n = A,T,C or G
US-09-606-421B-23

Alignment Scores:
Pred. No.:      8,32e-33      Length:      669
Score:          383.50       Matches:      83
Percent Similarity: 66.24%   Conservative: 21
Best Local Similarity: 52.87% Mismatches: 40
Query Match:    14.36%      Indels:       13
DB:             3           Gaps:           3

US-10-618-941-99 (1-507) x US-09-606-421B-23 (1-669)
Qy      349 AAlaValIleuAlaGluLeuProArg--ProArgArgProProLeuGlnTrpArgTyrSer 367
Db      7   GCCGTACTGCTGAATATCCCTGCAGACACGAAAGAACCAAGTTCAACATTGTACTCT 66
Qy      368 GluValSerPheMetGluLeuAryPheLeuGluAryValAryGanGlyTleTyrPro 387
Db      67 CAGTCACCAAGCTCTGAATTAATTAATTCCTGAAGATGTCAGGAATGGGATCTATCT 126
Qy      388 LeuMetAnPheAlaAlaThrArgProLeuGlyLeuProArgValIleuAlaProProPro 407
Db      127 CTGACAGCCCTT-----GGGCTGGCTCGGCCCCCAGCCAGCCACAGCAG 168
Qy      408 GluGluVal-----GlnuYsaIalYsThrProThrProGluProPhe 421
Db      169 GAGAGAGGACATCACTGTCTGTCGCCCTCTGTCAAGACTCCGACACCTGAACCACT 228
Qy      422 AspSerGluThrArgLysValIleGlnMetGlnCyaaLeuGluAryGSerGluAryLys 441
Db      229 GAGGTGAGACTCGCAAGGTGGTCTGATGCACTGAACATTGAAGTCGATGGAAGAGGA 288
Qy      442 AlaArgTrpHisLeuThrLeuLeuValLeuGluAryAryGLeuHisArgGlnLeuThr 461
Db      289 GTCAAAACCAACCTGACATTCTGTGAAGTGAAGGACAACTGAACCGGACCTGAGC 348
Qy      462 TyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuValHisTyrGly 481
Db      349 TGTGACTGATGCCAAATGAAGATATCCCGAGTTGGCGGCTGAGCTGTGACGTGGGC 408
Qy      482 PheLeuHisGluAryAspAryGMetLysLeuAlaIaPheLeuGlnSerThr 498
Db      409 TTCAATTAGTAGGCTGACCAAGACCGGTTGACTTCTGTGCTAAGAAAGACT 459
```

Search completed: January 14, 2006, 00:50:21
Job time : 275 secs

THIS PAGE BLANK (USPTO)

[illegible]

THIS PAGE BLANK (USPTO)

; Sequence 99, Application US/10618941
; GENERAL INFORMATION:
; APPLICANT: WHITE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 99
; LENGTH: 507
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-618-941-99
MAPAPAPRAREREREDSEDESIDLESFQGRWQKREOVNQGMPQLOSTPLAMDTREGV
EYWNELHFDGRKAFNAHEKIQTVPEQLVVDHPNIVKJHKYWLDTSEACARVIFITEYSSGS
LQPLKTKKNNHAKAMARAKRWCTQLLSALSPHACSPPI IHNLTSDTIFIQNGLIKISVW
HRIFSNALPPTALPDLSPIAESEELNHLFPPEYGEVADGTAVDIFSGMCALEMAVLEI
OTNGDTRVTEEAJARARHSLSDPNMREFLICCLADPARPSASHLSLFRHVLFEVHSLKLLAAHC
FIOHOYLMPENVVEEKTAMDIAVALELPRRPPLOMRYSVSEFWELDKPLEADVNGIYPLMN
PAATRLGLPRVLAAPPEEVQAKTTPPEPDSSTRKYIOMCNLERSEDKARHNTLLVLVEDR
LHRQLTYDLPTDSADLASELVHYGFLHEDDRKLAFLESTFLKTYGTQAI

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2006, 11:16:54 ; Search time 75 Seconds

(without alignments)
4769.371 Million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670

Sequence: 1 MAAPBAPRABRABRERERED.....RMKLAFLSTFLKRGCTQA 507

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1562.5	58.5	547	2	08BL77_MOUSE
2	1541	57.7	526	2	06GL31_XENTR
3	1537.5	57.6	535	2	099J45_MOUSE
4	1536	57.5	294	2	08NCX8_HUMAN
5	1534.5	57.5	535	2	09UHY1_HUMAN
6	1534.5	57.5	535	2	04R8X0_MACFA
7	1534.5	57.5	535	2	05RHH9_PONPY
8	1529.5	57.3	535	2	096SU3_HUMAN
9	1527.5	57.2	535	2	053F25_HUMAN
10	1519	56.9	532	2	05ZHL8_CHICK
11	1514	56.7	526	2	06NTZ7_XENLA
12	1486.5	55.7	519	2	05SNQ3_BRARE
13	1400.5	52.5	501	2	04SIG5_TETNG
14	1346	50.4	258	2	09NSY0_HUMAN
15	1331.5	49.9	483	2	04RKPS_TETNG
16	1283	48.0	258	2	091V36_MOUSE
17	1174.5	44.0	637	2	09Y0V6_DROME
18	1170	43.8	293	2	08WY55_HUMAN
19	1160	43.4	484	2	04T5T2_TETNG
20	1145.5	42.9	425	2	07Q3R5_ANOGA
21	965	35.6	324	2	0597G7_HUMAN
22	951.5	35.1	649	2	045668_CAEBL
23	778	29.1	157	2	08R3M0_MOUSE
24	395	14.8	2193	2	05VWF2_HUMAN
25	395	14.8	2297	2	07PUA4_ANOGA
26	392	14.7	379	2	056V74_MOUSE
27	378	14.2	848	2	056V74_MOUSE
28	378	14.2	1222	1	WNK4_MOUSE
29	378	14.2	1222	2	04VAC1_MOUSE
30	378	14.2	1677	2	081127_CAEBL
31	378	14.2	1743	1	WNK3_HUMAN

32	378	14.2	1743	2	05JRC1_HUMAN	05JRC1 homo sapien
33	378	14.2	1790	2	05JRC2_HUMAN	05JRC2 homo sapien
34	378	14.2	1800	2	08TCX6_HUMAN	08TCX6 homo sapien
35	378	14.2	1800	2	06UP76_HUMAN	06UP76 homo sapien
36	378	14.2	1838	2	018657_CAEBL	018657 caenorhabdi
37	377.5	14.1	438	2	053PL3_ORYSA	053PL3 oryza sativ
38	376	14.1	439	2	06DCU2_XENLA	06DCU2 xenopus lae
39	372	13.9	516	2	09STK6_ARATH	09STK6 arabidopsis
40	372	13.9	1222	1	WNK4_RAT	07PK6 rattus norv
41	370.5	13.9	1243	1	WNK4_HUMAN	096J92 homo sapien
42	370	13.9	1666	2	060QJ1_CAEBR	060QJ1 caenorhabdi
43	368.5	13.8	568	2	08S8Y9_ARATH	08S8Y9 arabidopsis
44	366	13.7	578	2	09LUV8_ARATH	09LUV8 arabidopsis
45	363.5	13.6	600	2	08H4Y8_ORYSA	08H4Y8 oryza sativ

ALIGNMENTS

RESULT 1
08BL77_MOUSE
ID 08BL77_MOUSE PRELIMINARY; PRT; 547 AA.
AC 08BL77;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length
DE enriched library, clone:B2034417 product:SIMILAR TO NUCLEAR RECEPTOR
DE BINDING PROTEIN (HLS7-INTERACTING PROTEIN KINASE) homolog.
GN Name=Nrbp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Eularchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RA MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."; Meth. Enzymol. 303:19-44(1999).
RL [2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RA MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Kono H., Kasukawa T., Saito R.,
RA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batelov S., Casavant H.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochava T.,
RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pebole G., Quackenbush J.,
RA Schirral L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Orido T., Furuno M., Hono H., Baldarelli R., Barin G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazarrelli J., Mombauts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohatsu S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).
RL [3]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RA THE PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:11617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama Y., Nishi K., Katsunai T., Taenito H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Ozawa Y., Izawa M., Ohara E., Matsubara M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384 format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RA Adachi U., Aizawa K., Akimura T., Arikawa T., Bono H., Carninci P.,
RA Fukuda S., Furuo M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirao T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ono M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takeki-Akaiwa S., Takeda Y., Tanaka T.,
RA Tomaru A., Tova T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
EMBL AK046142; BAC32612.1; -; mRNA.
DR Ensembl; ENSMUSG00000029148; Mus musculus.
DR MGI; MGI:2183436; Nrdp.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR Pfam; PF00069; Pkinase; 1.
DR Prodom; PD000001; Prot kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR ATP-binding; Kinase; Nucleotide-binding; Receptor; Transferase.
SQ SEQUENCE 547 AA; 61221 MW; 57462EA936E2762 CRC64;
Query Match 58.5%; Score 1562.5; DB 2; Length 547;
Best Local Similarity 62.6%; Pred. No. 2.7e-104;
Matches 310; Conservative 60; Mismatches 104; Indels 21; Gaps 6;
QY EREDESEDSIDLESPCGRWQKRRQVNGMNPGLQSTFLANDTEGEVVENNELHFGD 76
DB EEESESESESELESPCGRWQKRRQVNGMNPGLQSTFLANDTEGEVVENNELHFGD 103
QY 77 RKAFPAHEKIQVFEQVLVDHPNIVKHKYLDLSEACARVIFTEVSSGLKQFLK 136
DB RKAFPAHEKIQVFEQVLVDHPNIVKHKYLDLSEACARVIFTEVSSGLKQFLK 163
QY 104 RKQYKIQEERKVRVFNLLIQLEHINIVKFKYVADWENKARVIFTEYWSGSLKQFLK 196
DB 104 RKQYKIQEERKVRVFNLLIQLEHINIVKFKYVADWENKARVIFTEYWSGSLKQFLK 223
QY 137 KTKGNKAMNARAKMKCTQIISALSCSPPIIHGNTSTFTIHNGLIKISVWH 196
DB KTKGNKAMNARAKMKCTQIISALSCSPPIIHGNTSTFTIHNGLIKISVWH 223
QY 164 KTKGNKAMNARAKMKCTQIISALSCSPPIIHGNTSTFTIHNGLIKISVWH 255
DB 164 KTKGNKAMNARAKMKCTQIISALSCSPPIIHGNTSTFTIHNGLIKISVWH 277
QY 197 RIFSNALRPETALPDDLRSPIRAERBELNIAHFPPEYGEVAD-GTAVDIFSGMCALEM 255
DB 197 RIFSNALRPETALPDDLRSPIRAERBELNIAHFPPEYGEVAD-GTAVDIFSGMCALEM 277
QY 224 RIFAN-----VAPDTINNVKTCRBEQKNLHFAFBYGEVNVTTAVDIFYSGMCALEM 277
DB 224 RIFAN-----VAPDTINNVKTCRBEQKNLHFAFBYGEVNVTTAVDIFYSGMCALEM 277

QY 256 AVLEIOTNGDTR-VTEBIAIARAHSLSDPMNEPILCCARPARPASHSLFHRVLP 314
DB AVLEIOTNGDTR-VTEBIAIARAHSLSDPMNEPILCCARPARPASHSLFHRVLP 337
QY 315 VHSILKLAHACFIOHQYLPENNVVEEKTAMDLHVAIETPR-PRPPQWRSEVSFMS 373
DB VHSILKLAHACFIOHQYLPENNVVEEKTAMDLHVAIETPR-PRPPQWRSEVSFMS 397
QY 338 VESIKLAAHACFIOHQYLPENNVVEEKTAMDLHVAIETPR-PRPPQWRSEVSFMS 427
DB VESIKLAAHACFIOHQYLPENNVVEEKTAMDLHVAIETPR-PRPPQWRSEVSFMS 451
QY 374 LDKPLEDVNNGIYPLNFAATPRLGIPRYLADPPEEV-----QKATPTPEFDSERTR 427
DB LDKPLEDVNNGIYPLNFAATPRLGIPRYLADPPEEV-----QKATPTPEFDSERTR 451
QY 428 VYOMQCNERSDPAKRNHTLLVLEDRLHROLTDLPTDPAQDLASLVHYGFLEHD 487
DB VYOMQCNERSDPAKRNHTLLVLEDRLHROLTDLPTDPAQDLASLVHYGFLEHD 511
QY 452 VVLMQCNERSDPAKRNHTLLVLEDRLHROLTDLPTDPAQDLASLVHYGFLEHD 511
DB VVLMQCNERSDPAKRNHTLLVLEDRLHROLTDLPTDPAQDLASLVHYGFLEHD 511
QY 488 RKMLAAFLSTFLKY 502
DB RKMLAAFLSTFLKY 526
QY 512 QSRISVLEETLANKF 526
DB QSRISVLEETLANKF 526
RESULT 2
Q6GLJ1_XENTR PRELIMINARY; PRT; 526 AA.
ID Q6GLJ1_XENTR PRELIMINARY; PRT; 526 AA.
AC Q6GLJ1_XENTR PRELIMINARY; PRT; 526 AA.
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Nuclear receptor binding protein.
GN Name=nrp-prov.
OC Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnaa.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schaller G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepietkin M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywicki M.T., Skalska U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
EMBL BC074684; AAH74684.1; -; mRNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_thr_kinase.

RP NUCLEOTIDE SEQUENCE.
 RC TISSUE-Brain;
 RG The German cDNA Consortium;
 RA Ottemwælder B., Obermaier B., Deutschenhaur S., Schaiipp A.,
 RL Mewes H.W., Weill B., Amlid C., Oeangner A., Fodor G., Han M., Wiemann S.;
 RN Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE-Lung;
 RX MEDLINE=22368257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strauberg R.L., Feltingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hejeh F.,
 RA Datchenko L., Marusika K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedlin T.B., Tothilyuk S., Carninci P., Prange C.,
 RA Rha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smallue D.E.,
 RA Schnerch A., Schein J.B., Jones S.J.M., Matra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN NUCLEOTIDE SEQUENCE.
 RC TISSUE-Lung;
 RA Strauberg R.;
 RN Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN NUCLEOTIDE SEQUENCE.
 RA Sun H., Kozlowski A., Dignan G.;
 RT "The sequence of Homo sapiens BAC clone RP11-413M20.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN NUCLEOTIDE SEQUENCE.
 RA Waterston R.H.;
 RN Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN NUCLEOTIDE SEQUENCE.
 RA Waterston R.;
 RN Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN NUCLEOTIDE SEQUENCE.
 RA Wilson R.K.;
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF113249; AF21967.1; -; mRNA.
 DR EMBL; AK001946; BAA91993.1; -; mRNA.
 DR EMBL; AL136682; CAB66617.1; -; mRNA.
 DR EMBL; BC001221; AA01221.1; -; mRNA.
 DR EMBL; AC074117; AY14847.1; -; Genomic DNA.
 DR Ensembl; ENSG00000115216; Homo sapiens.
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0005070; F:SH3/SH2 adaptor activity; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.
 KM ATP-binding; Hypothetical protein; Kinase; Nuclear protein;
 KM Nucleotide-binding; Receptor; Serine/threonine-protein kinase;
 SO TRANSFERASE; 535 AA; 59845 MW; 398078661547EDD0 CRC64;

QY 17 EREDESESDILEESPGRWOKRRBOVNGNMPLGOSTFLAMDTSEGVVNNELHFGD 76
 DB 44 EEESESESESEIEEBCPGMOKREBEVNOBVPGLDSAYLANDTBEGVVNNVQFSE 103
 QY 77 RKAFAAHEEKIQVFQGLVLDHPNIVKLRWLDTSSEACARVIFITEVSSGLQFLK 136
 DB 104 RKNYKIQEEKVRVAFVFDLQLEHNTIVKFKYVADIKENKARVIFITEVSSGLQFLK 163
 QY 137 KTKGNKAMKAMKAMKWCCTOILSALSFACSPPIIHGNTISPTIFIOHGLIKGSVMH 136
 DB 164 KTKGNKAMKAMKAMKWCCTOILSALSYHSCDPIIHGNTICPTIFIOHGLIKGSV-- 221
 QY 197 RIFSNMLRPTALPDRLRPIRAREBELRLHFPPEYGVAD-GRVADIFSGMCALM 255
 DB 222 -----APDTINNVKTCREEQKMLHFAPEYGVTVTTAVADIVSGMCALM 269
 QY 256 AVLEIQTNQDTR-VTEBAIARASHLSDPNMRFEIICLARDPARPSAHLFHRVLPFE 314
 DB 270 AVLEIQNGSSVVPQEAISAIQLLEDPIQREFIQKIQSEPARRTARELLFHPALFE 329
 QY 315 VSHLKILAAHCFIQHOTLPENNVVEEKTAMDIAVLAELPR-PRRPLQWRYSVSEFME 373
 DB 330 VPSLKILAAHCFIQHOTLPENNVVEEKTAMDIAVLAELPR-PRRPLQWRYSVSEFME 389
 QY 374 LDKFLDVRNGIYPLNMFATRPLGLPRVLAPEPEV-----QKAKTPPEPDSSTRK 427
 DB 390 LDKFLDVRNGIYPLNMFATRPLGLPRVLAPEPEV-----QKAKTPPEPDSSTRK 443
 QY 428 VIOMQNLERSBDKARWHLTLVLVEDRLRQLYTYVLDPTDSAQDLASELVHYGFLHEDD 487
 DB 444 VVLMQNLERSBDKARWHLTLVLVEDRLRQLYTYVLDPTDSAQDLASELVHYGFLHEDD 503
 QY 488 RMTLAAFLSESTFLKY 502
 DB 504 QSRITSLLEETLTKPF 518

RESULT 6
 Q4R8X0_MACPA
 ID Q4R8X0_MACPA PRELIMINARY; PRT; 535 AA.
 AC Q4R8X0;
 DT 13-SEP-2005 (T-EMBLrel. 31, last sequence update)
 DT 13-SEP-2005 (T-EMBLrel. 31, last annotation update)
 DE Testis cDNA clone: Q4R8-11263, similar to human nuclear receptor
 binding protein (NRBP), (Brain cDNA, clone: Q4R8-17237, similar to
 human nuclear receptor binding protein (NRBP)).
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 OC Cercopithecoidea; Cercopithecoidea; Macaca.
 NC NCBI_TaxID=9541;
 RN NUCLEOTIDE SEQUENCE.
 RA International consortium for macaque cDNA sequencing, analysis;
 RT "DNA sequences of macaque genes expressed in brain or testis and its
 RT evolutionary implications.";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RN NUCLEOTIDE SEQUENCE.
 RA Orada N., Hitzata M., Tanuma R., Kueuda J., Hida M., Suzuki Y.,
 RA Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
 RT "Substitution rate and structural divergence of 5'UTR evolution:
 RT comparative analysis between human and cynomolgus monkey cDNAs.";
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB168327; BAB0451.1; -; mRNA.
 DR EMBL; AB169835; BAB01916.1; -; mRNA.
 DR RECEPTOR.
 SO SEQUENCE; 535 AA; 59845 MW; 398078661547EDD0 CRC64;

Query Match 57.5%; Score 1534.5; DB 2; Length 535;
 Best Local Similarity 61.6%; Pred. No. 2.7e-102;
 Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6;

DR Ensembl; ENSG00000115216; Homo sapiens.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR00719; Prot_Kinase.
 DR Pfam; PF00069; Kinase; 1.
 DR ProDom; PD000001; Prot_Kinase; 1.
 DR PROSITE; PS0011; PROTEIN KINASE DOM; 1.
 DR ATP-binding; Kinase; Nucleotide-Binding.
 KW Serine/threonine-protein kinase; transferase.
 KW SEQUENCE 535 AA; 59850 MW; 4829CA7CB60F0175 CRC64;
 SO Query Match 57.3%; Score 1529.5; DB 2; Length 535;
 Best Local Similarity 61.4%; Pred. No. 6.3e-102;
 Matches 304; Conservative 59; Mismatches 103; Indels 29; Gaps 6;
 QY 17 EREDESDSDILESPGKQKREOVNQGMPGLSTFLAMDTEEGEVVNNELHFGD 76
 D 44 EEEESDESEILBESPCGKQKREOVNQGMPGLSTFLAMDTEEGEVVNNELHFGD 103
 QY 77 RKAFAAEKIQVFEQVLVDHPIYKAKYMLDTSACARVIFTEYVSSGLKQFLK 136
 D 104 RKNYKQEEKRAVFDNLIOLEHLNIVKFKYMDIKENKARVIFITEYVSSGLKQFLK 163
 QY 137 KTKKHKAMAKKAKCTOILSALSFHACSPPIHGNLTSDTIFIOHNGLIKIGSV 196
 D 164 KTKKHKAMAKKAKCTOILSALSYHSCDPIIHGNLTCDTIFIOHNGLIKIGSV-- 221
 QY 197 RIFSNALRPPTALPDDRSPIRAERBELRMHFPPEYGEVAD-GTAVDIFSGMCALM 255
 D 222 -----ADDTINNHYKTCREBQKHLHFPAPYGEVNTTAVDVISFGMCALM 269
 QY 256 AVLEIQNGDTR-VTEBAIARHSLSDPNKRFEFLCCLARDPARRESASLLFHRVLF 314
 D 270 AVLEIQNGSSYVQPAISAIQLLEDPIQREFIOKLOSEPARRPTARELLFHPALFE 329
 QY 315 VHSIKLLAAHCFIOHGYLMPENVVEEKTAMDHLAVLAELPR-PRRPLQMRYSVSEFME 373
 D 330 VPSIKLLAAHCFIOHGYLMPENVVEEKTAMDHLAVLAELPR-PRRPLQMRYSVSEFME 389
 QY 374 LDKFLDVRNGIYPLMFAATRPGLPRVLAAPPEEV-----OKATPPEPDSSTRK 427
 D 390 LDKFLDVRNGIYPLTAF-----GLPRPQPOQEEVTSVPVPSVKTTPPEPAVESTRK 443
 QY 428 VIQMCNLERSEDKARHMLTLVLVEDRLRQLTVDLPDLSAODLASLAVHGFLEDD 487
 D 444 VVLMQCNIESEVGVKHLTLTLKLEDKLSRHLSCDLMPENIPELAAELVQGFISEAD 503
 QY 488 RMKLAFLSTFLKY 502
 D 504 QSRLTSLBETLNKF 518
 DB
 RESULT 9
 053F25 HUMAN PRELIMINARY; PRT; 535 AA.
 AC 053F25;
 DT 13-SEP-2005 (Tremblrel. 31, Created)
 DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
 DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
 DE Nuclear receptor binding protein variant (fragment).
 GN Name=NRBP;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC Tissue=Kidney;
 RA Maruyama K., Sugano S.,
 RT "Oligo-capping: a simple method to replace the cap structure of

RT eucaryotic mRNAs with oligoribonucleotides.";
 RL Gene 138:171-174 (1994).
 RL [2]
 RP NUCLEOTIDE SEQUENCE.
 RC Tissue=Kidney;
 RA Suzuki Y., Yoshitomo K., Maruyama K., Sugano A., Sugano S.,
 RT "Construction and characterization of a full length-enriched and a 5'-
 RT end-enriched cDNA library."
 RL Gene 200:149-156 (1997).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC Tissue=Kidney;
 RA Suzuki Y., Sugano S., Torok Y., Toyoda A., Takeda T., Sakaki Y.,
 RA Tanaka A., Yokoyama S.,
 RL Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK223136; BAD96856.1; -; mRNA.
 DR HGNC; HGNC:7993; NRBP.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 KW Receptor.
 KW NON TER
 SO SEQUENCE 535 AA; 59745 MW; 7FD909B54C33E89 CRC64;
 Query Match 57.2%; Score 1527.5; DB 2; Length 535;
 Best Local Similarity 61.4%; Pred. No. 8.8e-102;
 Matches 304; Conservative 59; Mismatches 103; Indels 29; Gaps 6;
 QY 17 EREDESDSDILESPGKQKREOVNQGMPGLSTFLAMDTEEGEVVNNELHFGD 76
 D 44 EEEESDESEILBESPCGKQKREOVNQGMPGLSTFLAMDTEEGEVVNNELHFGD 103
 QY 77 RKAFAAEKIQVFEQVLVDHPIYKAKYMLDTSACARVIFTEYVSSGLKQFLK 136
 D 104 RKNYKQEEKRAVFDNLIOLEHLNIVKFKYMDIKENKARVIFITEYVSSGLKQFLK 163
 QY 137 KTKKHKAMAKKAKCTOILSALSFHACSPPIHGNLTSDTIFIOHNGLIKIGSV 196
 D 164 KTKKHKAMAKKAKCTOILSALSYHSCDPIIHGNLTCDTIFIOHNGLIKIGSV-- 221
 QY 197 RIFSNALRPPTALPDDRSPIRAERBELRMHFPPEYGEVAD-GTAVDIFSGMCALM 255
 D 222 -----ADDTINNHYKTCREBQKHLHFPAPYGEVNTTAVDVISFGMCALM 269
 QY 256 AVLEIQNGDTR-VTEBAIARHSLSDPNKRFEFLCCLARDPARRESASLLFHRVLF 314
 D 270 AVLEIQNGSSYVQPAISAIQLLEDPIQREFIOKLOSEPARRPTARELLFHPALFE 329
 QY 315 VHSIKLLAAHCFIOHGYLMPENVVEEKTAMDHLAVLAELPR-PRRPLQMRYSVSEFME 373
 D 330 VPSIKLLAAHCFIOHGYLMPENVVEEKTAMDHLAVLAELPR-PRRPLQMRYSVSEFME 389
 QY 374 LDKFLDVRNGIYPLMFAATRPGLPRVLAAPPEEV-----OKATPPEPDSSTRK 427
 D 390 LDKFLDVRNGIYPLTAF-----GLPRPQPOQEEVTSVPVPSVKTTPPEPAVESTRK 443
 QY 428 VIQMCNLERSEDKARHMLTLVLVEDRLRQLTVDLPDLSAODLASLAVHGFLEDD 487
 D 444 VVLMQCNIESEVGVKHLTLTLKLEDKLNRHLSCDLMPENIPELAAELVQGFISEAD 503
 QY 488 RMKLAFLSTFLKY 502
 D 504 QSRLTSLBETLNKF 518
 DB
 RESULT 10
 05ZHL8 CHICK PRELIMINARY; PRT; 532 AA.
 AC 05ZHL8;
 DT 25-OCT-2004 (Tremblrel. 28, Created)
 DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)

```

DE      25-OCT-2004 (TREMBLrel. 28, last annotation update)
DE      Hypothetical protein.
GN      ORNames=RCMH04_3518;
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC      Gallus.
CX      NCBI_TaxId=9031;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=CB; TISSUE=Bursa;
RA      Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
RA      Fiedler P., Kutter S., Blagoderzki A., Kostowska D., Koter M.,
RA      Plachy J., Caminci P., Hayashizaki Y., Buerstedt J.M.;
RT      "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT      genefunction analysis.";
RL      Genome Biol. 6:R6-R6(2005).
DR      EMBL; AJ721116; CAG32775.1; -; mRNA.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR      GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR      InterPro; IPR000719; Prot_kinase.
DR      InterPro; IPR002290; Ser_thr_kinase.
DR      Pfam; PF00069; Pkinase; 1.
DR      Prodom; PD000001; Prot_kinase; 1.
DR      SMART; SM00220; S_TKc.1.
DR      SMART; SM00219; TYRc.1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM, 1.
DR      Hypothetical protein.
KW      SEQUENCE 532 AA; 59693 MW; 2176262D10AEC76 CRC64;

```

Query Match	56.9%	Score 151.9;	DB 2;	Length 532;
Beet Local Similarity	59.6%;	Pred. No. 3.6e-101;		
Matches	300;	Conservative	63;	Mismatches 124; Indels 16; Gaps 3
QY	6	PAPRAREREREDSESDILES	PCGRNOKREQVNOGMPLOSTFLAMDTBEGV	65
DB	33	PYTSAAPEEEEESEDES	ETLSESPCGNRQKREEVNOQNVGIDSAIYAMDTBEGV	92
QY	66	EYVNNELHFGDKARPAHEEKIQTVEQOLVLDHPNIVKLAHKYMLDTSEACARVIFTEY		125
DB	93	EYVNNELVQFSEKKNFQLQEEEEKAVFDNLQLEHNIVKFKYKMDVKNKARVIFTEY		152
QY	126	VSSGSLKQFLKTKTKNKNKAMNARKMWTQISLSPHAAQSPRIHENTLSDTIFIOH		185
DB	153	MSSGSLKQFLKTKTKNKNKMTNEMAKRMWTQISLSTSYHSCDPIIHENLTCDTIFIOH		212
QY	186	NGLIIGSYWHRIFSNALRPPLPDDLSPIRAEEREIRNLHFFPEYGEVAD-GTAVD		244
DB	213	NGLIIGSV-----ADPTINNHYKTCREBQKULHFAPEYGEVAAVTTAVD		258
QY	245	IFSFGMCALMAVLEIQNGDTR-VTEEAIAPARHSLSDPNKREFILCCLADPARBPA		303
DB	259	IYSPGMCALMAVLEIQNGESSYVQGEAINGAIQLLEBPLOREVIOKLEBDDPGKRPYA		318
QY	304	HSLLHRRVLYFEVHSLKTLAAHCFIQOYIMPEVNVBEKTKAMDIAVLAELRPPRPPIQ		363
DB	319	RELLHRRQALFEVPSKTLAAHCTIVGQHMI PENALAEEMTKNIDMSVLAIEINHADRBEVX		378
QY	364	WRYSEVSFMELEKPLEDVANGIYPLMNFATRPPLGRLPRLAPRPEEVOAKPTPTDEPPDS		423
DB	379	MIFSOSPALIEKLFEDVANGIYPLTAFGMPPRPQPOQVAVKSPILAPEVVKPTTDEPAEV		438
QY	424	ETRKATIQONCNDERSEDKARMHLLTLVLVDEDRILHQITVLDLPTDSAOCLASGLVYGPL		483
DB	439	ETRKAVLMOQCNIESVEEGVGHLLTLVLKDEKLTNHLSDCLDPNDNIOGLAEVLOLGI		498
QY	484	HEDDRMKLAFLLESTFLKTRGTQ		506
DB	499	SEADQSRUTCLEEAFFSKFYVTR		521

RESULT 11			
06NT27_XENLA	PRELIMINARY	PRR	526 AA.
ID	06NT27_XENLA	PRELIMINARY	PRR
AC	06NT27		
DT	05-JUL-2004 (Tremblrel. 27, Created)		
DT	05-JUL-2004 (Tremblrel. 27, Last sequence update)		
DT	05-JUL-2004 (Tremblrel. 27, Last annotation update)		
DE	MG881395 protein.		
GN	Name=MG881395		
OS	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;		
OC	Xenopodinae; Xenopus; Xenopus.		
OX	NCBI_TaxId=8355;		
OX	1]		
RP	NCILOTIDE SEQUENCE.		
RC	TIS5B=Embryo;		
RC	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins P.S., Wagner L.H., Shenmen C.F., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.M., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Scapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein J.U., Ushed T.B., Toshiyuki S., Carninci P., Penge C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,		
RA	Bosak S.L., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Pehey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Blakesley A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Butterfield J.S.N., Krzywinski M.I., Skalska U., Smallie D.E.,		
RA	Schnecker A., Schein J.E., Jones S.J.M., Marra M.A.,		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RL	1]		
RP	NCILOTIDE SEQUENCE.		
RC	TIS5B=Embryo;		
RC	MEDLINE=22311132; PubMed=12454917; DOI=10.1002/dvdy.10174;		
RA	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,		
RA	Richardson P.,		
RT	"Genetic and genomic tools for Xenopus research: The NIH Xenopus		
RT	Initiative."		
RL	Dev. Dyn. 225:384-391(2002).		
RL	1]		
RP	NCILOTIDE SEQUENCE.		
RC	TIS5B=Embryo;		
RC	Klein S., Strausberg R.,		
RL	Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.		
RL	EMBL, BC068805; AA068805.1; -; mRNA.		
DR	GO: GO:0005524; F:ATP binding; IEA.		
DR	GO: GO:0004674; P:protein-tyrosine/threonine kinase activity; IEA.		
DR	GO: GO:0004113; P:protein-serine/threonine kinase activity; IEA.		
DR	GO: GO:0016740; P:transferase activity; IEA.		
DR	GO: GO:0051301; P:cell division; IEA.		
DR	GO: GO:0004668; P:protein amino acid phosphorylation; IEA.		
DR	InterPro: IPR007019; Prot kinase.		
DR	InterPro: IPR022290; Ser Thr kinase.		
DR	InterPro: IPR01245; Tyr kinase.		
DR	Pfam: PF00069; Pkinase; 1.		
DR	Prodom: PD000001; Prot kinase; 1.		
DR	SMART: SM00220; S_TKc; 1.		
DR	SMART: SM00219; TYKc; 1.		
DR	SMART: SM00219; TYKc; 1.		
DR	PROSITE: PS0011; PROTEIN KINASE DOM; 1.		
DR	SEQUENCE 526 AA; 5936 MW; E966FC5FCBC7428 CRC64;		

Query Match	56.7%;	Score 1514;	DB 2;	Length 526;
Best Local Similarity	58.8%;	Pred. No. 8.1e-101;		

SQ SEQUENCE 519 AA; 58644 MM; CDPFD99609BDC18C4 CRC64;
 Query Match 55.7%; Score 1486.5; DB 2; Length 519;
 Best Local Similarity 58.7%; Pred. No. 7,7e-99;
 Matches 297; Conservative 66; Mismatches 106; Indels 37; Gaps 7;

 QY EPARRRREREREDESDSDLESPPCGRMQKREOVNQGMPLGOSTFLAMDTEEG 64
 :
 Db QPAP-----SATEDEESESESESLIESPCGRMQKRREEVQNRPVIDNAYIAMDTEG 84
 29 QPAP-----SATEDEESESESESLIESPCGRMQKRREEVQNRPVIDNAYIAMDTEG 84

 QY VEVVVNNELHPDQRKAFAAHEEKITVEQLVDHPIVKLHKWMLTSEACARVIFITE 124
 :
 Db VEVVNNENVMFSERKNFKLQEBKAAVNDNLQLEHLIVTFHKWADYKENRAVVIITE 144
 85 VEVVNNENVMFSERKNFKLQEBKAAVNDNLQLEHLIVTFHKWADYKENRAVVIITE 144

 QY YVSSGSLSKPFLKTKKKHKKANARAWRCWTQILSALSFHACSPPIIHGNLTSDTIPIQ 184
 :
 Db YMSGSGSKLPFLKTKKKHKKANARAWRCWTQILSALSYHSCEPPIIHGNLTCDTIPIQ 204
 145 YMSGSGSKLPFLKTKKKHKKANARAWRCWTQILSALSYHSCEPPIIHGNLTCDTIPIQ 204

 QY HNGLIKTIGSWHRFNSNALRPTPLPDLDSPFAESELRLHFPEPEGEVAD-GTAV 243
 :
 Db HNGLIKIKIGSV-----APTINNHWTCREEOKSLFFAFBEYGAVNAVTTAV 250
 205 HNGLIKIKIGSV-----APTINNHWTCREEOKSLFFAFBEYGAVNAVTTAV 250

 QY DIFSFGCALEMAVLIEIQTGDDTR-VYEBAIARARSLSDPNNREFLLCCIARPARPS 302
 :
 Db DIVSFGCALEMAVLIEIQSGESSYVSQEAINSALQSLEDPLQEFLOKCLEADVPSKPT 310
 251 DIVSFGCALEMAVLIEIQSGESSYVSQEAINSALQSLEDPLQEFLOKCLEADVPSKPT 310

 QY AHSLLFRVLFEVASLKLTLAAHCFIOHOYMPENVVEKTKAMDLHAVALPRPRRPPL 362
 :
 Db ABELLFPQALFEVYQKLTLAAHCIVSHQHMI PENALAEITRKMDPNQVIME-----RKEN 365
 311 ABELLFPQALFEVYQKLTLAAHCIVSHQHMI PENALAEITRKMDPNQVIME-----RKEN 365

 QY QMRYSBVSFMELDKFLEDVRNGIYPLMNPATRPGLPRVLAAPPPEEQ-----KAKTP 416
 :
 Db QLKKSQEPALBLDFLEDRVNRGIYPLTAFA-----GMPCTQQOQAQVAKSPIVPBSYKTP 419
 366 QLKKSQEPALBLDFLEDRVNRGIYPLTAFA-----GMPCTQQOQAQVAKSPIVPBSYKTP 419

 QY TPEPDSERTKVOMOCNLRSESDKAWMH.TLLVLVEDRLRHQTLYOLPTDSAQDIASE 476
 :
 Db TPPEABETRTKVVOQCNIIEPBVGAGHHLLTLKLBKLNRLHSCDLAPNEQBIAIVE 479
 420 TPPEABETRTKVVOQCNIIEPBVGAGHHLLTLKLBKLNRLHSCDLAPNEQBIAIVE 479

 QY LVHYGFLEHDDRMTAAFLSTFLPKY 502
 :
 Db LVQJGFISBGDQPRLTIVLEAPERF 505
 477 LVHYGFLEHDDRMTAAFLSTFLPKY 502
 480 LVQJGFISBGDQPRLTIVLEAPERF 505

RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 RT the early vertebrate proto-karyotype." ;
 RL Nature 431:946-957(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC -1- FUNCTION: Plays a key role in the control of the eukaryotic cell
 CC cycle. It is required in higher cells for entry into S-phase and
 CC mitosis. Component of the kinase complex that phosphorylates the
 CC repetitive C-terminus of RNA polymerase II. Catalytic component of
 CC MPF (By similarity).
 CC -1- SUBUNIT: Forms a stable but non-covalent complex with cyclin B in
 CC mature oocytes (By similarity).
 CC EMBL; CAE01014768; CAG05517.1; -; Genomic_DNA.
 CC InterPro: IPR000719; Prot_kinase.
 CC InterPro: IPR002290; Ser_Thr_kinase.
 CC InterPro: IPR001245; Tyr_kinase.
 CC Pfam: PF00669; Pkinase; 1.
 CC ProDom: PD000001; Prot_kinase; 1.
 CC SMART: SM00220; S_TKc; 1.
 CC SMART: SM00219; TYKc; 1.
 CC PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
 FT NON_TER 1 501
 FT TER 1 501
 SQ SEQUENCE 501 AA; 57651 MW; C28F26487BA61567 CRC64;

Query Match 52.5%; Score 1400.5; DB 2; Length 501;
 Best Local Similarity 55.3%; Pred. No. 1.2e-92;
 Matches 280; Conservative 56; Mismatches 83; Indels 87; Gaps 5;

43 QVNGNMPGQISTFLAMDEEGVYVNNELHFDGRKAFANHEKIQTEVQVLVDHPNI 102
 DB 1 QVNGNMPGQISTFLAMDEEGVYVNNELHFDGRKAFANHEKIQTEVQVLVDHPNI 60
 QY 103 VRLHKWLDTESEACARVIFTEYVSSGSLKQFLKTKKKNKANAAMKWCQIIISAL- 161
 DB 61 VRLHKWLDTESEACARVIFTEYVSSGSLKQFLKTKKKNKANAAMKWCQIIISAL- 120
 QY 162 -----SFLHACSPRIHGNLSDPTIFIOHNLKIGSVNHR 197
 DB 121 VSLILYIMRFLMTSPDDLFPQSYIHSCDPIIHGNLCTDTIFIOHNLKIGSVNHR 180
 QY 198 IFSNALRPTALPD-DLRSEIRAREBELRNLAFFPEYGEVA--DGTAVDIFSGMCALE 254
 DB 181 LFVN-----VFPDASVHGKGRORHDEQRNLHFFAPYGGSGEDDYAIDIFSGI CALE 234
 QY 255 MATLEIQTNGDTVTBEALIRARHSLSDPNMR----- 286
 DB 235 MAVLEIQANGDSVYSKALVNAQHSLEDPMRYSVNRTPCRYTTSQGWLYITLLTRL 294
 QY 287 -----EFILCCLARPARPSASHLIFHRYLFEVHSLKLL 321
 DB 295 CSLISCNWSPETEGCNFTLDFLOEFTQSCLRHNAKLRPTAADLIFHRVLFVHSLKLL 354
 QY 322 AAHCFIQHOYLMBENVVEEKTAMDIAVLAELPRPRRPLQWRYSVEVSFMEIDKFLLEV 381
 DB 355 AAHCLINNOYLLENCEEEKTKSIDPAVVAELIRHNDROGVOLAKYHVSLELDKFLLEV 414
 QY 382 RNCIYVLNMFARPLGLPRVLAAPPEEVOKATPTPEPDSSTRKYIQCNCLESSEDK 441
 DB 415 KNGIYVLNMFASMPHPVPRLALSQEOYETVKTPTPEPETRKYVQHCNLESNEEG 474
 QY 442 ARWHLTLVLLEDRILRQLTYVDLLPT 467
 DB 475 TKTHLSLFLKMDDKLHNLSCDIFPS 500

RESULT 14
 Q9NSYO_HUMAN

ID Q9NSYO_HUMAN PRELIMINARY; PRT; 258 AA.
 AC Q9NSYO;
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
 DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
 DE Hypothetical protein DKFZp434P086 (LOC340371 protein).
 GN Name=DKFZp434P086; Synonyms=LOC340371;
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RA Koehler K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Ebert L., Schlick M., Neubert P., Schatten R., Henze S., Korn B.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL137662; CAB70864.1; -; mRNA.
 DR EMBL; CR457350; CAG33631.1; -; mRNA.
 DR PIR; T46491; T46491.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro: IPR000719; Prot_kinase.
 DR ProDom: PD000001; Prot_kinase; 1.
 KW Hypothetical protein; Nucleotide-binding.
 SQ SEQUENCE 258 AA; 29852 MW; 3C866AAD4CDE26EA CRC64;

Query Match 50.4%; Score 1346; DB 2; Length 258;
 Best Local Similarity 100.0%; Pred. No. 4.4e-89;
 Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

250 MCALEAVALEIQNGDTRTVEALIRARHSLSDPNMRFFILCCLARPARPSASHLTP 309
 DB 1 MCALEAVALEIQNGDTRTVEALIRARHSLSDPNMRFFILCCLARPARPSASHLTP 60
 QY 310 RVLFEVHSLKLLAAHCFIQHOYLMBENVVEEKTAMDIAVLAELPRPRRPLQWRYSVEY 369
 DB 61 RVLFEVHSLKLLAAHCFIQHOYLMBENVVEEKTAMDIAVLAELPRPRRPLQWRYSVEY 120
 QY 370 SFMEIDKFLLEVDRNGIYPLNMFARPLGLPRVLAAPPEEVOKATPTPEPDSSTRKYI 429
 DB 121 SFMEIDKFLLEVDRNGIYPLNMFARPLGLPRVLAAPPEEVOKATPTPEPDSSTRKYI 180
 QY 430 OMOCNLESSEDKARWHLTLVLLEDRILRQLTYVDLLPTQSADDLASELVHGYGLHEDDM 489
 DB 181 OMOCNLESSEDKARWHLTLVLLEDRILRQLTYVDLLPTQSADDLASELVHGYGLHEDDM 240
 QY 490 KLAAFLESTFLKYRGTOA 507
 DB 241 KLAAFLESTFLKYRGTOA 258

RESULT 15
 Q9ARKP5_TETNG
 ID Q9ARKP5_TETNG PRELIMINARY; PRT; 483 AA.
 AC Q9ARKP5;
 DT 13-SEP-2005 (T-EMBLrel. 31, Created)
 DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
 DE Chromosome 5 SCAR15026, whole genome shotgun sequence.
 DE (Fragment).
 GN ORName=GSTENG00032841001;
 OS Tetraodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 OC NCBI_TaxID=99883;

```

RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jellison O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Mucelli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Desliiva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Authouard V., Jubin C., Castell V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Cactolico L., Poulsen J., De Bernardis V.,
RA Cuand C., Duprat S., Broctier P., Couanceau J.P., Gouy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volf J.N., Guigo R., Zody M.C., Mesirov J.,
RA Landrad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Landier V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weisenbach J., Roest Crolius H.,
RT "genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope: Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -!- FUNCTION: Plays a key role in the control of the eukaryotic cell
CC cycle. It is required in higher cells for entry into S-phase and
CC mitosis. Component of the kinase complex that phosphorylates the
CC repetitive C-terminus of RNA polymerase II. Catalytic component of
CC MPF (By similarity).
CC -!- SUBUNIT: Forms a stable but non-covalent complex with cyclin B in
CC mature oocytes (By similarity).
CC EMBL: CA601015026; CAG11037.1; -, Genomic_DNA.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR Pfam: PF00069; Kinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TYKc; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
FT NON_TER 1
FT NON_TER 483
SQ SEQUENCE 483 AA; 54608 MW; 1A970E4D3EFA224 CRC64;

Query Match 49.9%; Score 1331.5; DB 2; Length 483;
Best Local Similarity 53.6%; Pred. No. 1,1e-87;
Matches 265; Conservative 70; Mismatches 114; Indels 45; Gaps 4;

QY 19 EDESEDESDILESPGCGKQKRRQVQGNMPLQSTFLAMPTEEGVGVVWNLHPGDRK 78
DB 8 DEESDDESEILESPGCGKQKRRQVQGNMPLQSTFLAMPTEEGVGVVWNLHPGDRK 67
QY 79 AFAAHEKIQTVFEQVLVDHPNIVKLKTYMLDTSEACARVIFTEYVSSGSLKQFLKKT 138
DB 68 NFKQLEKVKAVVDNLIHLHANIVKFKYKWDTKGKRAVIFITEYVSSGSLKQFLKKT 127
QY 139 KKHAKAMNARAKMKTQIISAL-----SFLHACSPPII 172
DB 128 KKHAKMNERKALKMKTQIISALKTLSVQALQDAVKSILMFSPFCYLIHSDPPII 187
QY 173 HGNLTSTFIQHNGLIKIGVWHIRIFSNLRRPTALPDDLRSPIRAEREELNTHFPFP 232
DB 188 HGNLTCTFIQHNGLIKIGSV-----APDTINNHVKTCYEEQKYLHFPYAP 233
QY 233 EYGEVADGTAVDIFSPGMCALEMAVALIEIQTNQDTR-VTEBAIARARSHLSDPMNRFPILC 291
DB 234 EYGDNDNTTAVDIFSPGMCVLEMAVLEIHNGSSSYSGDAINNAIQLLEDPLQKELIQK 293
QY 292 CLARDPARPSANSLFHRVLFYVHSLKLLAACHFCIOHYLMPENVVEEYKTKAMDHLAVL 351
DB 294 CLSDSDPEVRPTAREBLFDPALFEVPLKLLAASHIVHQYIMDENALEEMTKNLDPLVI 353
QY 352 ABLPRPRRPPLQWRYSVSEVMELDKFLQEDVRNGIYPLMNPATRPGLPRVLAPPEEVQ 411

```

```

DB 354 AEM-----KEDIQKLQGFPALELDKFLQEDVRNGIYPTATGCLSPQKQGVVKSPTVVV 409
QY 412 KAKTPPEPDSFTRKYIQNCNLERSEDKARWHLTLVLVEDRLHRLTYDLLPTDSAQ 471
DB 410 LVSPPTPEPALELRRVQWQCNIELIEEGTKFHLTLKLLEDKLANHLLSCDMLPNESVQ 469
QY 472 DLASELVHGFILH 485
DB 470 ELAGELVLELISB 483

```

Search completed: January 12, 2006, 11:21:58
Job time : 78 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2006, 11:22:06 ; Search time 115 Seconds

(without alignments)
1842.082 Million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670

Sequence: 1 MAAPBPAPRRARERERED.....RMKLAFLSTFLKNGTQA 507

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA Main:

1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep:*

2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep:*

3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep:*

4: /cgn2_6/prodata/1/pubppaa/US10A_PUBCOMB.pep:*

5: /cgn2_6/prodata/1/pubppaa/US10B_PUBCOMB.pep:*

6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2670	100.0	507	4	US-10-618-941-99
2	2499	93.6	439	5	US-10-840-512-185
3	1580	59.2	326	3	US-09-862-027-10
4	1580	59.2	326	5	US-10-989-328-10
5	1534.5	57.5	335	3	US-09-840-787-20
6	1534.5	57.5	335	4	US-10-153-668-322
7	1534.5	57.5	335	3	US-09-764-875-749
8	1529.5	57.3	335	4	US-10-059-585-2
9	1508.5	56.5	335	4	US-10-040-647-10
10	1444	54.1	501	3	US-09-925-297-609
11	1310	49.1	434	4	US-10-153-668-434
12	1174.5	44.0	637	6	US-11-097-143-270
13	946	35.4	340	3	US-09-764-868-798
14	946	35.4	340	3	US-09-764-875-1053
15	882.5	33.1	732	4	US-10-292-798-1714
16	882.5	33.1	886	4	US-10-017-161-2068
17	756	28.3	302	3	US-09-764-868-822
18	756	28.3	302	3	US-09-764-875-750
19	543	20.3	231	5	US-09-862-027-36
20	543	20.3	231	5	US-10-989-328-36
21	415.5	15.6	461	3	US-09-862-027-35
22	415.5	15.6	461	5	US-10-989-328-35
23	395	14.8	779	4	US-10-353-529-49
24	395	14.8	1345	4	US-10-433-794-17
25	395	14.8	2193	5	US-10-490-592-5
26	395	14.8	2245	4	US-10-618-941-108
27	392	14.7	2132	5	US-10-840-512-195

ALIGNMENTS

28	386	14.5	1751	5	US-10-840-512-197	Sequence 197, App
29	378	14.2	1069	4	US-10-182-743-50	Sequence 50, App1
30	378	14.2	1743	4	US-10-460-545-2	Sequence 2, App1
31	376.5	14.1	1234	4	US-10-052-648A-36	Sequence 36, App1
32	372	13.9	516	3	US-09-862-027-34	Sequence 34, App1
33	372	13.9	516	5	US-10-989-328-34	Sequence 8, App1
34	371.5	13.9	1243	4	US-10-478-146-8	Sequence 37, App1
35	370.5	13.9	1231	4	US-10-052-648A-37	Sequence 4, App1
36	370.5	13.9	1231	4	US-10-196-935A-4	Sequence 40, App1
37	368.5	13.8	1601	3	US-09-862-027-40	Sequence 40, App1
38	368.5	13.8	1601	3	US-10-989-328-40	Sequence 12, App1
39	367.5	13.8	557	4	US-10-052-648A-12	Sequence 80, App1
40	367.5	13.8	1251	4	US-10-114-270-80	Sequence 170860, App1
41	365.5	13.7	658	4	US-10-424-599-170860	Sequence 72414, A
42	365.5	13.7	746	4	US-10-425-114-72414	Sequence 103929, App1
43	363.5	13.6	601	4	US-10-437-963-103929	Sequence 38, App1
44	359.5	13.5	670	4	US-10-052-648A-38	Sequence 2, App1
45	359.5	13.5	2108	5	US-10-491-467-2	

RESULT 1

US-10-618-941-99

Sequence 99, Application US/10618941

Publication NO. US2004019792A1

GENERAL INFORMATION:

APPLICANT: MANTHE, DAVID

APPLICANT: MANTHE, DAVID

APPLICANT: CAENEPEL, SEAN

TITLE OF INVENTION: NOVEL KINASES

FILE REFERENCE: 034536-0321

CURRENT APPLICATION NUMBER: US/10/618,941

PRIOR FILING DATE: 2003-07-15

PRIOR APPLICATION NUMBER: 60/395,632

NUMBER OF SEQ ID NOS: 143

SOFTWARE: PatentIn version 3.2

SEQ ID NO 99

LENGTH: 507

TYPE: PRT

ORGANISM: Homo sapiens

US-10-618-941-99

Query Match 100.0%; Score 2670; DB 4; Length 507;
Best Local Similarity 100.0%; Pred. No. 1.4e-220;
Matches 507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAAPBPAPRRAREREREDSESDIIEBSPGGRWQRRQVQGNMPGLQSTFLAND 60
DB	1	MAAPBPAPRRAREREREREDSESDIIEBSPGGRWQRRQVQGNMPGLQSTFLAND 60
QY	61	TEBGEVWVNEHFGGRKAPFAHEEKTQVFPQVLYVHDPNIVKHLKYLDTSEACARYI 120
DB	61	TEBGEVWVNEHFGGRKAPFAHEEKTQVFPQVLYVHDPNIVKHLKYLDTSEACARYI 120
QY	121	FTTEVYSSGSLKQFLKTKKNKAMARAKWCQIOLISALFLHACSPPIHGNLTSDT 180
DB	121	FTTEVYSSGSLKQFLKTKKNKAMARAKWCQIOLISALFLHACSPPIHGNLTSDT 180
QY	181	FTIHNGLIKIGSVHRIFSNLRPPALPDDLRSPIRAERELRNLFPPPEYGEVADG 240
DB	181	FTIHNGLIKIGSVHRIFSNLRPPALPDDLRSPIRAERELRNLFPPPEYGEVADG 240
QY	241	TAVDIFSPGMCALMAVVEIOTNGDRTVTEEAIAARHSLSPNNREFITCCLADPAR 300
DB	241	TAVDIFSPGMCALMAVVEIOTNGDRTVTEEAIAARHSLSPNNREFITCCLADPAR 300
QY	301	PSAHSILFRVLFVYVSLKLAHCFIQQOYIMPEVVEKTKAMDIAVLAELPRPRRP 360
DB	301	PSAHSILFRVLFVYVSLKLAHCFIQQOYIMPEVVEKTKAMDIAVLAELPRPRRP 360

Qy	361	PLQMRYSSEFSFMEIDFLEBEVRNGIPLANNPATRPLGIPRLAPPEEVOKATTPPER	420
Db	361	PLQMRYSSEFSFMEIDFLEBEVRNGIPLANNPATRPLGIPRLAPPEEVOKATTPPER	420
Qy	361	PLQMRYSSEFSFMEIDFLEBEVRNGIPLANNPATRPLGIPRLAPPEEVOKATTPPER	420
Db	361	PLQMRYSSEFSFMEIDFLEBEVRNGIPLANNPATRPLGIPRLAPPEEVOKATTPPER	420
Qy	421	FDSESTRVIVMOONLRSRSEDKAAMHLLTLLVLEDRILHQLTYDLLPDSADLASSEVHY	480
Db	421	FDSESTRVIVMOONLRSRSEDKAAMHLLTLLVLEDRILHQLTYDLLPDSADLASSEVHY	480
Qy	421	FDSESTRVIVMOONLRSRSEDKAAMHLLTLLVLEDRILHQLTYDLLPDSADLASSEVHY	480
Db	421	FDSESTRVIVMOONLRSRSEDKAAMHLLTLLVLEDRILHQLTYDLLPDSADLASSEVHY	480
Qy	481	GFLEHEDDRMKLAAPLESTFLKRGTOA	507
Db	481	GFLEHEDDRMKLAAPLESTFLKRGTOA	507

```

US-10-840-512-185
2
RESULT 2
1
Sequence 185, Application US/10840512
2
Publication No. US20050125852A1
3
GENERAL INFORMATION:
4
APPLICANT: CAENEPEEL, SEAN
5
APPLICANT: MANNING, GERRARD
6
APPLICANT: CHARYDCZAK, GLEN
7
APPLICANT: CRICORIEV, IGOR
8
TITLE OF INVENTION: NOVEL KINASES
9
FILE REFERENCE: 034536-1455
10
CURRENT APPLICATION NUMBER: US/10/840,512
11
CURRENT FILING DATE: 2004-05-07
12
PRIOR APPLICATION NUMBER: 60/469,014
13
PRIOR FILING DATE: 2003-05-09
14
NUMBER OF SEQ ID NOS: 239
15
SOFTWARE: PatentIn version 3.2
16
SEQ ID NO 185
17
LENGTH: 499
18
TYPE: prt
19
ORGANISM: Mus musculus
20
US-10-840-512-185

```

Query Match	93.6%	Score 2499	DB 5	Length 499
Best Local Similarity	93.9%	Pred. No. 6.8e-206		
Matches 476; Conservative	11	Mismatches 12	Indels 8	Gaps 2

Oy	MAAPEPRARAREKEREDESEDEDIIEBSPCORMKREBOVNOGMPLOSTFLAMD	60
Db	1 MAAPERPARG--REBEREDSEDESDIIEBSPCORMKREBOVNOGMPBIOSTFLAMD	58
Oy	TEBSEVYVNNELHFGRKAPAAHEEKIQTVPEQVLVDHPNI VKLHKWLDTSEACARVI	120
Db	59 TEEBSEVYVNNELHFGRKAPAAHEEKIQTMFEQLLVHPNIVKLHKWLDASEARAVI	118
Oy	PIFEYVSSGLKQFLKTKTKKHKAMABAKMKCTQIISALSFJLHACSPPIIHENLISDT	180
Db	119 PIFEBYSSGLKQFLKTKTKKHKAMABAKMKCTQIISALSFJLHACSPPIIHENLISDT	178
Oy	181 IFIQHNGLIKIGSVWHRIFSNALRPETALPDDLRSPIABREBELRNLIHFPEYGEVADG	240
Db	179 IFIQHNLIKIGSVWYRIFSN-----ALPDLRSPIABREBELRNLIHFPEYGEVADG	232
Oy	241 TAYDIBSEFGCALEMVLEIQTNGDTRVYEEAIAABRSHSDPMNKEIILCCLLADPARR	300
Db	233 TAYDIBSEFGCALEMVLEIQAGDTRVYEEAIAABRSHSDPMNKEIILCCLLADPARR	292
Oy	301 PSAHSLLFHRVLEFEVSHKLKLAHCPIQHQYIMPENVEBEETKAMDILAVLAELPRPRP	360
Db	293 PSAHNLIFHRVLEFEVSHKLKLAHCPIQHQYIMPENVEBEETKAMDILAVLAELPRQHPG	352
Oy	361 PLQWRSEVSEFMELDKFLBDVRNGIYPLMNPATPBLGPRVLAPPEBEOKATPTPEP	420
Db	353 PMQWRSEVSEFLDKFLBDVRNGIYPLMNPFAAARPLGIPRVLAPPEEOKATPTPEP	412
Oy	421 FDESTRVIVIMOCNLRSESDKARWHLTLLVVEDRLHOLRYDILLPTPSAODLASLIVHY	480
Db	413 FDESTRVIVIMOCNLRSESDKARWHLTLLVVEDRLHOLRYDILLPTPSAODLASLIVHY	472
Oy	481 GFLHEDDRMKLAFLSTFLKTRGTQA	507

Db 473 GFLHEDDRTKLAFLFTTLKRYGTQA 499

RESULT 3
 US-09-862-027-10
 Sequence 10 Application US/09862027
 Patent No. US20020142428A1
 GENERAL INFORMATION:
 APPLICANT: Hodge, Martin R.
 TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof
 FILE REFERENCE: 35800/234862
 CURRENT APPLICATION NUMBER: US/09/862,027
 CURRENT FILING DATE: 2001-05-21
 PRIOR APPLICATION NUMBER: US 09/345,473
 PRIOR FILING DATE: 1999-06-30
 NUMBER OF SEQ ID NOS: 82
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 10
 LENGTH: 326
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: VARIANT
 LOCATION: (1)...(326)
 OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-862-027-10

Query March	59.2%	Score 1580;	DB 3;	Length 326;
Best Local Similarity	97.4%	Pred. No. 4e-127;		
Matches 301; Conservative	0;	Mismatches 2;	Indels 6;	Gaps 1.

```

QY      1  MAAPBAPBAREREREREDSESDSLEESPCGRNOKREOVNOQNMGLOSTPLAMD 60
Db      12  MAAPBAPBAREREREREDSESDSLEESPCGRNOKREOVNOQNMGLOSTPLAMD 71
QY      61  TEEGVEVWVNEHFPDRAPAAHEEKIQTVESQVLVDHPNIVYLAHKYMLDTSEACARVY 120
Db      72  TEEGVEVWVWNEHFPDRAPAAHEEKIQTVESQVLVDHPNIVYLAHKYMLDTSEACARVY 131
QY      121  FITEYVSSGSLKQPLKTKKHKAMNAPAKRMCTQILSALSFHACSPPIHGULTSDT 180
Db      132  FITEYVSSGSLKQPLKTKKHKAMNAPAKRMCTQILSALSFHACSPPIHGULTSDT 191
QY      181  IFQHNGLIKIGSVWHRFSNALRPPALPDLLSPRARERIRNHFPPPEYGEVADG 240
Db      192  IFQHNGLIKIGSVWHRFSN-----ALPDLLSPRARERIRNHFPPPEYGEVADG 245
QY      241  TAVDIFPFCMALENAVLEIQNGDTRVTEBAIARAHSLSDPMREFFILCCLARPAR 300
Db      246  TAVDIFPFCMALENAVLEIQNGDTRVTEBAIARAHSLSDPMREFFILCCLARPAR 305
QY      301  PSAHSLFFH 309
Db      306  PSAHSLFFH 314

RESULT 4
US-10-989-228-10
/ Sequence 10, Application US/10989228
/ Publication No. US20050089917A1
/ GENERAL INFORMATION:
/ APPLICANT: Hodge, Martin R.
/ TITLE OF INVENTION: Novel Kinases and Uses Thereof
/ FILE REFERENCE: 35800/234862
/ CURRENT APPLICATION NUMBER: US/10/989, 228
/ CURRENT FILING DATE: 2004-11-15
/ PRIOR APPLICATION NUMBER: US/09/862,027
/ PRIOR FILING DATE: 2001-05-21
/ PRIOR APPLICATION NUMBER: US 09/345,473
/ PRIOR FILING DATE: 1999-06-30
/ NUMBER OF SEQ ID NOS: 82
SOFTWARE: FastSeq for Windows Version 4.0

```

```

1  SEQ ID NO 10
2  LENGTH: 326
3  TYPE: prt
4  ORGANISM: Homo sapiens
5  FEATURE:
6  NAME/KEY: VARIANT
7  LOCATION: (1)...(326)
8  OTHER INFORMATION: Xaa = Any Amino Acid
9  US-10-989-228-10
10
11 Query Match          59.2%; Score 1580; DB 5; Length 326;
12 Best Local Similarity 97.4%; Pred. No. 4e-127;
13 Matches 301; Conservative 0; Mismatches 2; Indels 6; Gaps 1;
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
```

```

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <unknown>
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 535 amino acids
TYPR: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAUTTOT3
CLONE: 864683
SEQUENCE DESCRIPTION: SEQ ID NO: 20 :
US-09-840-787-20

Query Match          57.5%, Score 1534.5; DB 3; Length 535;
Best Local Similarity, 61.6%; Pred. No. 6.5e-123;
Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6,

OY      17 EREDSEDESDIIEESPCGRWQKREBOVNOGMMFGLOSTFIAMDTEBGVEVMNDELHFD 76
DB      44 EEESESEDESIIIEESPGRWQKREBOVNOGVGIDSAYLAMDTEBGVEVMNEVQFSR 103
OY      77 RKAFAAHEKIQTYPEDQLVDNHNIYKLHKYWLDTSEACARVFTIREYSSGSLLKQFLK 136
DB     104 RKNYKLOBEKRAVPDMILQLEHLNTIVFKHYMDIKENKRVFTIREYSSGSLLKQFLK 163
OY     137 KTKGNKKAMANAARAKRWCTOLLSALSFHACSPTIHGNLTSPITFOHNGLIKIGSVWH 196
DB     164 KTKGNKHTMEKAKRKCTQLSLSTLHSCDPIIHGNLTCDTIFIOHNGLIKIGSV-- 221
OY     197 RIFSNAIRPPTALRPDDLRSPIRAERBELRLHFPEPYGEVAD-GTAVDIISFGMCALEM 255
DB     222 -----ABDTINNHVKTCREBQKNLHFAPRYEGEVTNVTTAVDIYSFCMCALEM 269
OY     256 AVLEIQTNNGDR-VTEBAIARARSLSDDPMKEFITLCGLADPARRSASLSLPHRVLF 314
DB     270 AVLBIQNGBSYYPQEAISSAIQLDLBDPLQREFIQKLOSSEPARRPTABELFHPPALFE 329
OY     315 VHSJLKLAACHFIOHOVLMPENVVEBKTKAMDIAVAELPFR-PRRPPLQWRYSVSFME 373
DB     330 VPSJLKLAACHFIOHMIPBNALBEETKNNDSAVLAELIPAGCGRPVQVLNSQSPLAE 389
OY     374 LDKLELDVRNGIYELMNFAATRPGLRVLAPPEEV-----OKAKTPPEPPDSSTRK 427
DB     390 LDKLELDVRNGIYELTFN-----GLPRPQQPQEEVTSVPVPSVKTTPPERPALEVETR 443
OY     428 VIQMOCULESSEDARAHMLTLVLLELRRLRLQTLTDLPTDSAADLASLVHGFLHEDD 487
DB     444 VLMQCNIESIESEBVKHILLTLKLBDKLNRLSLCDLMPMBNIPETAALVELVOLGISPAD 503
OY     488 RMKLAFLAESSTFLKY 502
DB     504 QSRLLSLLEETLNKF 518

RESULT 6
US-10-153-668-322
Sequence 322, Application US/10153668
Publication No. US20030092616A1
GENERAL INFORMATION:
APPLICANT: HONDA, Goichi
APPLICANT: MATSUDA, Akio
APPLICANT: MURAMATSU, Shuji
APPLICANT: ISHIZAWA, Kenya
TITLE OF INVENTION: STATE Activating Gene
FILE REFERENCE: 1254-0207P
CURRENT APPLICATION NUMBER: US/10/153,668

```



```

/ PRIOR APPLICATION NUMBER: PCT/JPO00/05060
/ PRIOR FILING DATE: 2000-07-28
/ PRIOR APPLICATION NUMBER: US 60/183,322
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: US 60/159,590
/ PRIOR FILING DATE: 1999-10-18
/ PRIOR APPLICATION NUMBER: JP 2000-118776
/ PRIOR FILING DATE: 2000-01-11
/ PRIOR APPLICATION NUMBER: JP 2000-183767
/ PRIOR FILING DATE: 2000-05-02
/ PRIOR APPLICATION NUMBER: JP 11-248036
/ PRIOR FILING DATE: 1999-07-29
/ NUMBER OF SEQ ID NOS: 64
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 535
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-059-585-2

```

Query Match	57.3%	Score 1529.5	DB 4	Length 535
Best Local Similarity	61.4%	Pred. No. 1.7e-122		
Matches 304, Conservative	59	Mismatches 103	Indels 29	Gaps 6

```

Qy 17 ERDESEDESDILIEESPCGRMQRGRQVQGMPPGLOSTFLAMDTBESEVVMNIEHGD 76
Db 44 EEBEESDESELEESPCGRMQRGRREVQGRVPGIDSAIYAMDTEBEVVMNIEVQOSE 102
Qy 77 RKAFAAHEEKIQVFEQOLVLDHPNIVKLAHKTWLDTSEACARVIFITBEVSGSLKOPFK 134
Db 104 RKNYKLOEBEVRVRFNPLQLBEHLNVKHKYMWADIKENKARVIFITBEVMSGSLKOPFK 163
Qy 137 KTKGNKAMAMARAKMKCTOILSALSFLACSPPIIHGNTLSDTFLFIQHGLIKIGSVH 196
Db 164 KTKGNKHTNNEKAMKMKCTOILSALSFLACSPPIIHGNTLSDTFLFIQHGLIKIGSV-- 222
Qy 197 RIFSMLRPPYLPDDLKSPPIAREBELNLHFPPBEVEVAD-CTAVDIFSFGKALEM 255
Db 222 -----APDPIINHVKTCRBEQNLHFAPBEYGEVNVVTAVDIFSFGKALEM 266
Qy 256 AVLEIQTNQDTR-VTEEAIRARHSISDPMMEFILLCCIARDPARPSAHSILFTRVLYFE 314
Db 270 AVLEIQTNQESSVYRPEALISSAQLLEDDLYQNEBFIQKLCQSPARRPFAKELHTRPALFE 322
Qy 315 VHSKLTLAAHCFIQOYLYMPENVEEKTAMDLAHVALBEP-PRRPELOMRYSSEVSEME 372
Db 330 VPSLTKLTLAAHCFVGHQHMIPENALBEITQNMDSAVLAIEIPAGPBEVPOVLTYGSPALTE 388
Qy 374 LDEFELEDVNGSIGVILNMPAATRPGLGPRYLARPREAV-----QKAKTPTEPPEPSERK 422
Db 390 LDEFELEDVNGSIGVILYAF-----GLPRQOPQOBEVTSVPVPSVKPTPEPRPAVEERK 443
Qy 428 VVICMGNLSESEDKARMLTLLVLEDRLAHQLTVDLLPTDSADOLASELVNHYGFLHEDD 487
Db 444 VVICMGNNISSEVGEGVGHITLLTKLEDKSRHLSCDLMPENIPETALAEVLQGFISBAD 503
Qy 488 RMLTAAFLSESTPKY 502
Db 504 QSRULTLEETLNF 518

```

RESULT 9
US-10-040-647-10
/ Sequence 10, Application US/10040647
/ Publication No. US20030092154A1
/ GENERAL INFORMATION:
/ APPLICANT: (US only) ANTALIS Tom Marie and HOOPER John David
/ TITLE OF INVENTION: NOVEL MOLECULES
/ NUMBER OF SEQUENCES: 30
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
/ STREET: 400 GARDEN CITY PLAZA
/ CITY: GARDEN CITY

STATE: NEW YORK
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/040,647
FILING DATE: 07-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/023,942
FILING DATE: <Unknown>
APPLICATION NUMBER: POS101/97
FILING DATE: 13-FEB-1997
APPLICATION NUMBER: PP0422/97
FILING DATE: 18-NOV-1997
APPLICATION NUMBER: International PCT Application
FILING DATE: 13-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: DIGIGLIO, FRANK S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 11168
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742 4343
TELEFAX: (516) 742 4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 535 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-040-647-10

		Query Match	56.5%; Score 1508.5; DB A 4; Length 535;
		Best Local Similarity	60.8%; Pred. No. 1,1e-120;
		Matches 301; Conservative	60; Mismatches 105; Indels 29; Gaps 6;
Qy	17	E R D E S D S D I L B E S F C G W Q K R R E O V N G N M P G L O S T I L A M D T E B G E V V M N E H F G D	76
Dz	44	E E E E S D S E B I I E E S P C G R W Q K R R E V N O R N P G I D S A Y I A M D T E G E V V N N E V O F S E	103
Qy	77	R K A P A A H E K I Q V F Y E D V L V D H P N Y I K A H K W M L D P S E A C A R Y I P T E Y V S G S L K Q F L K	136
Dz	104	R K Y I K L O E E K Y C A V F D U I O J E H N I N Y K F K I A D I K E N K A R Y I F I T G Y S S G S L K O F L K	163
Qy	137	K T K K N K A M N A R A M K R W C T O I L S A L S P L A C S P I I H G N I T S D T I F I O H N G I K I G S V M H	196
Dz	164	K T Q K N H Q T N N E K A M K R W C T O I L S A L S Y H S C D P I I H G N I T C D T I F I O H N G I K I G S V -	221
Qy	197	R I T S N M L R P P T A L P D D L R S P I R A R E R E L R Y L H F P P P Y G E V A D - G R A V D I F S T G M C L E M	255
Dz	222	-A P D I T N N H V K C R E Q K N L H P A P Y G S V N T V T A A V D I Y S G M C L A G M	269
Qy	256	A V L E I Q T N G D T R - V T E F A I R A R H S L S D P M R E F I I C C L A R D P A R R S A H S L F H R U L F E	314
Dz	270	A V E I Q G N G S S Y V P Q E A I S S A I Q L L E D P I Q R F I Q K C L O S F A R R A R T A E L L F H R A L F E	329
Qy	315	V H S I K L I A A C F I O H Y T L M E B N V V E E K T A M D L A V I A E L P R - P R R P P L O M R Y S E V S F M E	373
Dz	330	V P S I K L I A A C I G H O H M I P E N A L E E T K M D M S A V A E I P A G P G R E P V O T L I V S O S F A L Z E	389
Qy	374	L D F L E D V R N G I Y P L N M F A T R P L G I P R V I A P P E E V - - - - - Q X K T P P E Z F D S E T T K	427
Dz	390	L D F L E V R N G I Y F L T A F - - - - - G L P R Q Q P O E E V T S R V V P S Y K T P P E A E V E T R K	443
Qy	428	V I O M O C N L E S E D I C A R M H L T L L I V L E D R L H R O L T Y D L L P T D S A O D I A S E L V H Y G F L H E D	487
Dz	444	V I M O C N I E S E V G V K H N L T L L I K E D K A N R H I S C O L M P E N I P E T L A A E L V O G F T S E A D	503

QY 488 RMKLAFLSTFLKY 502
DB 504 QSRITSLTLEETLNKF 518

RESULT 10
US-09-925-297-609
; Sequence 609, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 609
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (29)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-609

Query Match 54.1%; Score 1444; DB 3; Length 501;
Best Local Similarity 57.9%; Pred. No. 3, 6e-115;
Matches 293; Conservative 64; Mismatches 117; Indels 32; Gaps 7;

QY 6 PPRAREREREDSEDESDILESPCGRWQKREOVNQGMPLOSTFLAMDTBEDEV 65
DB 2 PPRAREREREREDSEDESDILESPCGRWQKREOVNQGMPLOSTFLAMDTBEDEV 58
QY 66 EVVWNEIHFGRKAPAAHBEKIQTEVQVLVDHPNIVKLHKWMLDTSBACARVITFEY 125
DB 59 EVVWNEIHFGRKAPAAHBEKIQTEVQVLVDHPNIVKLHKWMLDTSBACARVITFEY 118
QY 126 VSSGSLKQFLKTKKQKAKANARAWKQCTQILSALSPHACSPRIHGNLTSDTIFIQH 185
DB 119 VSSGSLKQFLKTKKQKAKANARAWKQCTQILSALSPHACSPRIHGNLTSDTIFIQH 178
QY 186 NGLIKIGSVWHRFSAALAPRTALPDLSPIAREBEELRNHFFPEYGEVAD-GTAVD 244
DB 179 NGLIKIGSVWHRFSAALAPRTALPDLSPIAREBEELRNHFFPEYGEVAD-GTAVD 224
QY 245 IFSFGCALEMAVLEIQTNGDTR-VTEBAIARARHSLSDPNKREFTLCCIARPARPSA 303
DB 225 IFSFGCALEMAVLEIQTNGDTR-VTEBAIARARHSLSDPNKREFTLCCIARPARPSA 284
QY 304 HSLFLRVLPEVNSILKLLAHCPTIQOYLMPENVEBEKTKANDLAAVLAELPR-PRRPL 362
DB 285 HSLFLRVLPEVNSILKLLAHCPTIQOYLMPENVEBEKTKANDLAAVLAELPR-PRRPL 344
QY 363 QWRISVSVFELDKFLPEDEVNNGIYPLMNPATRPGLPRVLAAPPEV-----QAKATP 416
DB 345 QWRISVSVFELDKFLPEDEVNNGIYPLMNPATRPGLPRVLAAPPEV-----QAKATP 398
QY 417 TPEPDSERKVLQMOCNLSESEDKARWHLTLVLVEDRLHROTLYDLPTDQAODLASE 476
DB 399 TPEPDSERKVLQMOCNLSESEDKARWHLTLVLVEDRLHROTLYDLPTDQAODLASE 458
QY 477 LVHYGFLHEDDRMKLAFLSTFLKY 502
DB 477 LVHYGFLHEDDRMKLAFLSTFLKY 502

DB 459 LVQLGFISEADQSRITSLTLEETLNKF 484

RESULT 11
US-10-153-668-434
; Sequence 434, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MATSUDA, Akio
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAWA, Kenya
; TITLE OF INVENTION: STAT6 Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153,668
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260681
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 434
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-668-434

Query Match 49.1%; Score 1310; DB 4; Length 434;
Best Local Similarity 63.8%; Pred. No. 9, 8e-104;
Matches 263; Conservative 42; Mismatches 83; Indels 24; Gaps 6;

QY 17 EREDESEDESDILESPCGRWQKREOVNQGMPLOSTFLAMDTBEDEVNNEIHFED 76
DB 44 EREDESEDESDILESPCGRWQKREOVNQGMPLOSTFLAMDTBEDEVNNEIHFED 103
QY 77 RAAFAAHEKIQTEVQVLVDHPNIVKLHKWMLDTSBACARVITFEYVSSGSLKQFLK 136
DB 104 RAAFAAHEKIQTEVQVLVDHPNIVKLHKWMLDTSBACARVITFEYVSSGSLKQFLK 163
QY 137 KTKKQKAKANARAWKQCTQILSALSPHACSPRIHGNLTSDTIFIQHNGLIKIGSVW 196
DB 164 KTKKQKAKANARAWKQCTQILSALSPHACSPRIHGNLTSDTIFIQHNGLIKIGSVW 221
QY 197 RIFSAALAPRTALPDLSPIAREBEELRNHFFPEYGEVAD-GTAVDIFSFGCALEM 255
DB 222 RIFSAALAPRTALPDLSPIAREBEELRNHFFPEYGEVAD-GTAVDIFSFGCALEM 269
QY 256 AVLEIQTNGDTR-VTEBAIARARHSLSDPNKREFTLCCIARPARPSAHSILFHRVLE 314
DB 270 AVLEIQTNGDTR-VTEBAIARARHSLSDPNKREFTLCCIARPARPSAHSILFHRVLE 329
QY 315 VHSILKLLAHCPTIQOYLMPENVEBEKTKANDLAAVLAELPR-PRRPLQWRYSSEVFW 373
DB 330 VHSILKLLAHCPTIQOYLMPENVEBEKTKANDLAAVLAELPR-PRRPLQWRYSSEVFW 389
QY 374 LDKFLPEDEVNNGIYPLMNPATRPGLPRVLAAPPEVQAKATPTEBPDSER 425
DB 390 LDKFLPEDEVNNGIYPLMNPATRPGLPRVLAAPPEVQAKATPTEBPDSER 434
RESULT 12
US-11-097-143-270
; Sequence 270, Application US/11097143
; Publication No. US20050208558A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Venter, J. Craig
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
/ TITLE OF INVENTION: DROSOPHILA GENES.
/ FILE REFERENCE: C1000728
/ CURRENT APPLICATION NUMBER: US/11/097,143
/ CURRENT FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-19
/ PRIOR APPLICATION NUMBER: 60/161,932
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
/ NUMBER OF SEQ ID NOS: 43008
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 270
/ LENGTH: 637
/ TYPE: PRT
/ ORGANISM: DROSOPHILA
/ US-11-097-143-270

Query Match      44.0%; Score 1174.5; DB 6; Length 637;
Best Local Similarity 49.3%; Pred. No. 7.4e-92;
Matches 245; Conservative 77; Mismatches 146; Indels 29; Gaps 9;

16 REHEDESEDESDLESPGCRMKQKREBOVNGMPPGLQSTFLAMDPEGEVVMNHLHG 75
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
84 RESGDSEDESELESPGCRMLKREEDVDQRPGLDCHVHMLDPEGEVVMNVSQVA 143
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
76 DRKAPAAHEKIQTFVEQVLVDHPNIVKLHKYMLDTSEA-CARVIFITEYSSGSLKOF 134
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
144 SIQELKSQSEKQKQVVDNLQDHQNIKFKRYTDTQABRRRVVFTIEYSSGSLKOF 203
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
135 LKTKKNKAMNARAWKRCQIISLSFLHACSPPIHGNLTSDTIFIOHNGLIGSV 194
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
204 LKTKKNAKRLPLESWRMCTQIISLSFLHACSPPIHGNLTSDTIFIOHNGLIGSV 263
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
195 WHRIFENALRPALDDLRSPFRAERELRNHFF-PPRYGVADGT-AVDIFSGMCA 252
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
264 -----VPPAVHYSVRGRERERERERGAHYQAPPEYGAADLTALDIYAGMCA 313
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
253 LEMAVLEIQTNCG--DTRVTEEAJARARHSISDPNMEFFLCCLPARPPRPSASHLFFR 310
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
314 LEMAALEIQPSNESSTAINETIQRITFSELDLQDLRKCLANPQODRPSANDLLFFR 373
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
311 VLPEVHSLKLLAHGCTIOHQLMPENVVEKTKAMD-----LHAVLAELPRRRPPL 362
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
374 LTFEVHSLKLLTFHCLV---FSPANRTFSEFADGLMQRYQPDVVAQQLLAGGQR 429
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
363 OKRYSVSFME-LDKLEEDVRNGIYPLMPAATRPGLRVLAAPPEEYQAKTTPPEEP 421
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
430 QYRLAVDSGADKLEKFEVDKYGVPPLIYSGKKPNF--RSRAASPERADSVSATPEV 488
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
422 DSETRKVIQMCNLERSEDKARWHLTLVLVEDRLRQLTVDLPTDSADQLASSELVHG 481
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
489 DRESRIIVMMCSVKIKEDSNDITWTITLMDKMKRQITCVQNEPDADLTSELVRLG 548
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
482 FLHEDDRMKLAFLSEST 498
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
549 FVHLDDQDKIQVLLERT 565
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 13
US-09-764-868-798
/ Sequence 798, Application US/09764868
/ Patent No. US2002016871A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PT232
/ CURRENT APPLICATION NUMBER: US/09/764,868
/ CURRENT FILING DATE: 2001-01-17
/ Prior application data removed - refer to PALM or file wrapper
/ NUMBER OF SEQ ID NOS: 1510
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 798
/ LENGTH: 340
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (274)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (289)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (296)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (307)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (319)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-798

Query Match      35.4%; Score 946; DB 3; Length 340;
Best Local Similarity 60.4%; Pred. No. 1.4e-72;
Matches 201; Conservative 29; Mismatches 71; Indels 32; Gaps 7;

117 ARVIFITEYSSGSLKOPFKTKKNKANARAWKRCQIISLSFLHACSPPIHGNL 176
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5 ARVIFITEYSSGSLKOPFKTKKNKANARAWKRCQIISLSFLHACSPPIHGNL 64
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
177 TSDTIFIOHNGLIGSVVHRIFSNALRPPTALPDLSPIRAERELRNHFFPEYGE 236
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
65 TCDTIFIOHNGLIGSV-----APDTINNHVTCREQGNLFFFAEYGE 110
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 VAD-GTAVDIFSPGCMLEMAVLQTNCGDTR-VTEEARARHSISDPNMEFFLCCLA 294
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
111 VTNVTTAVDIYSGCMLEMAVLQTNCGESSYVDEAISAIOLEEDLQREFFLOKQ 170
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
295 RDPARPPRPSASHLFFRVHLEPVASHKLLAHGCTIOHQLMPENVVEKTKAMDHAVLAEL 354
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
171 SEPARRPTRELTFRALPEVPSKLLAHGCTIOHQLMPENVVEKTKAMDTSANLAI 230
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
355 PR-PRRPPQWRYSEVSFEMLDKFLQEDVNGIYPLMPAATRPGLRVLAAPPEEYQAK 413
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231 PAGPGEPEVQTLYSQSPALELDFLEEDVNGIYPLTAFA-----GLPRXQOQOBEVTS 284
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
414 KTPPEPDPSETRKVIQMCNLERSEDKARWHL 446
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
285 VVP-EXVXDFTXTSLRWL-----ARWXL 308
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 14
US-09-764-875-1053
/ Sequence 1053, Application US/09764875
/ Patent No. US20040018969A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: P202
```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2006, 11:22:31 ; Search time 31 Seconds

(without alignments)
154,622 Million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670

Sequence: 1 MAPBPAPRPRARERERED.....RMKLAFLSTFLKRGTOA 507

Scoring table:

BLOSUM62

Searched: 67062 seqs, 9454214 residues

Total number of hits satisfying chosen parameters: 67062

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Published Applications AA New:
1: /cgn2_6/ptcodata/1/pubppa/US06_NEW_PUB.pep:*
2: /cgn2_6/ptcodata/1/pubppa/US07_NEW_PUB.pep:*
3: /cgn2_6/ptcodata/1/pubppa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptcodata/1/pubppa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptcodata/1/pubppa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptcodata/1/pubppa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptcodata/1/pubppa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptcodata/1/pubppa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	199	7.5	588	6	US-10-821-234-1137
2	192	7.2	366	6	US-10-661-426-7
3	191	7.2	585	7	US-11-134-563-16
4	190	7.1	366	6	US-10-661-426-4
5	190	6.9	366	6	US-10-661-426-13
6	184.5	6.9	524	6	US-10-878-556A-131
7	184	6.9	591	6	US-10-770-726-71
8	170.5	6.4	348	6	US-10-661-426-12
9	168.5	6.3	403	6	US-10-523-477-11
10	168.5	6.3	403	6	US-10-770-726-83
11	168.5	6.3	403	6	US-10-770-726-85
12	168.5	6.3	403	7	US-11-109-156-12
13	168.5	6.3	403	7	US-11-092-168-1
14	167.5	6.3	348	6	US-10-661-426-9
15	167.5	6.3	348	6	US-10-661-426-14
16	163	6.1	256	6	US-11-113-424-183
17	157	5.9	1302	7	US-11-004-057-6
18	154	5.8	672	7	US-11-004-057-2
19	154	5.8	1493	7	US-11-004-057-4
20	153	5.7	277	7	US-11-151-601-4
21	151	5.7	1493	7	US-11-004-057-21
22	149	5.6	278	6	US-10-055-877-149
23	149	5.6	278	6	US-11-103-065-4
24	149	5.6	278	7	US-11-151-601-10
25	145	5.4	664	6	US-10-485-517-308

26	142	5.3	712	6	US-10-770-726-66	Sequence 66, App1
27	140.5	5.3	275	6	US-10-523-477-10	Sequence 10, App1
28	140.5	5.3	970	6	US-10-770-726-86	Sequence 86, App1
29	140	5.2	547	6	US-10-770-726-87	Sequence 87, App1
30	139.5	5.2	532	6	US-10-860-501-7	Sequence 7, App1
31	138.5	5.2	297	6	US-10-770-726-48	Sequence 48, App1
32	138.5	5.2	297	7	US-11-109-156-11	Sequence 11, App1
33	138	5.2	445	6	US-10-770-726-70	Sequence 70, App1
34	137	5.1	656	6	US-10-821-234-1121	Sequence 1121, App1
35	135.5	5.1	456	6	US-10-860-501-5	Sequence 5, App1
36	135.5	5.1	500	6	US-10-860-501-4	Sequence 4, App1
37	133.5	5.0	231	7	US-11-151-601-5	Sequence 5, App1
38	133.5	5.0	231	7	US-11-151-601-12	Sequence 12, App1
39	133.5	5.0	1311	6	US-10-509-422-5	Sequence 5, App1
40	131	4.9	258	6	US-10-770-726-51	Sequence 51, App1
41	130.5	4.9	418	7	US-11-099-691-1	Sequence 1, App1
42	129.5	4.9	418	7	US-11-109-156-2	Sequence 2, App1
43	129.5	4.9	1236	7	US-11-115-086-4	Sequence 4, App1
44	129.5	4.9	1306	6	US-10-995-561-905	Sequence 905, App1
45	127	4.8	1061	7	US-11-059-814-18	Sequence 18, App1

ALIGNMENTS

```
RESULT 1
US-10-821-234-1137
/ Sequence 1137, Application US/10821234
/ Publication No. US20050255114A1
/ GENERAL INFORMATION:
/ APPLICANT: Labat, Ivan
/ APPLICANT: Steache-Crain, Birgit
/ APPLICANT: Andarmant, Susan
/ APPLICANT: Tang, Y. Tom
/ TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
/ FILE REFERENCE: 821A
/ CURRENT APPLICATION NUMBER: US/10/821,234
/ CURRENT FILING DATE: 2004-04-07
/ PRIOR APPLICATION NUMBER: US 60/462,047
/ PRIOR FILING DATE: 2003-04-07
/ NUMBER OF SEQ ID NOS: 1704
/ SOFTWARE: pc_seq_genes Version 1.0
/ SEQ ID NO 1137
/ LENGTH: 588
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-821-234-1137

Query Match      7.5%, Score 199; DB 6; Length 588;
Best Local Similarity 20.6%; Pred. No. 7e-10;
Matches 125; Conservative 87; Mismatches 213; Indels 182; Gaps 27;

QY      5 EPAPR-----AREREREDSEDESD-----ILRESP 34
      |||||
DB      9 DPAFRPRFRGSGSETTGGAAAAAAGGSETRGSETSEVSEGGPRAAAVMSDSGA 68
      |||||

QY      35 GHWQKR-----EYVNOGMPGLQSTFLAMDTEGVEVYNELHFGDKAFAMBEKIQ 88
      |||||
DB      69 LPMISIRDDYELQEVIGSGATAVQAACAPKKE---KVAIKRINL-----EKCO 115
      |||||

QY      89 TYFEQV-----LVDHENIVKHKYMDTSEACARVFTREYVSGSLQKPLK--TK 139
      |||||
DB      116 TSNDELTKETQAMSGCHPHIVSYYSFVAKDE---LWLWKLISGSGVLDITKRIIVAK 171
      |||||

QY      140 KHKKA--NNARAKWCTQILSLSPFIHSCPTIIGNTLSDPTIFIOHGLIKIGVMHR 197
      |||||
DB      172 GHHKSGVDESLTATITAREVLGSLYTHKNGQ--IHRDVKAGNIIILGEGSVQIADFGVS 229
      |||||

QY      198 IFSNALRPPTALPDDIRSPIRAREELRNILH-----FPPPEYGEVADG--TAVDIFSPGM 250
      |||||
DB      230 AF-----LATGCDI-----TRNKVTRKTFVGTPCWMAPEVMEQVGRGYDFRADISFGSI 276
      |||||

QY      251 CALDMA-----VLEIQTNQGDTRVTEBAIARARHSLSDPM-----REPI 289
      |||||
```

```
Db 277 TAIELATGAAPYHAKYPPMKVLMTLQNDPPSLETGV-----QDKEMLKTKYKSPRKM 329
Qy 290 LCCIADDPARRPSAHSLLFHRVLFVYHSLKTLAAHCFIOHYLMPEVNVBEKTKAMD LHA 349
Db 330 SLCLOQDPKPRPFAELHHRKFPQAKNKE-----FLOEKTLORAPPTISERAK----- 377
Qy 350 VLAELPRPRPPLQ-----WRYSVSPMELDKFLVDVRNGIYPLAN----- 390
Db 378 -----KVRVPRSSGRKHKTEDEGWMWSDDER---DESEBCKAALISQLRSPRVESIS 428
Qy 391 ----FAATRPPLGIPRVLADPPEEV-----QXAKTPEPPPESETRKVIQ 430
Db 429 NSELFPPTDPVG---TLQVPEQISAMLPQAPQIATOPTQVSLPPTAEP--AKTQALIS 483
Qy 431 MOCNLERSEDKAMHMTLLVLVEDRLHRQUTYDLE--TSAQDLASELVHYGFLHEDDM 489
Db 484 SSGSGQ--ETKIPISIVLRIRNSKKELANDIRFEFTGRDTPAGVSGELISAGLVGRDLV 541
Qy 490 KLAAPLE 496
Db 542 IVANALQ 548

RESULT 2
US-10-661-426-7
; Sequence 7, Application US/10661426
; Publication No. US20050262584A1
; GENERAL INFORMATION:
; APPLICANT: Sneen, Jen
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Asai, Tsuneaki
; APPLICANT: Tena, Guillaume
; TITLE OF INVENTION: Master Activators of Pathogen Responsive
; TITLE OF INVENTION: Genes
; FILE REFERENCE: 00786/397003
; CURRENT APPLICATION NUMBER: US/10/661,426
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: PCT/US 02/07650
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/275,199
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-661-426-7

Query Match 7.2%; Score 192; DB 6; Length 366;
Best Local Similarity 22.6%; Pred. No. 1.5e-09;
Matches 82; Conservative 53; Mismatches 122; Indels 106; Gaps 16;

Qy 45 NOGNMGLQSTFLANDTEBGE-----VVMNELHFGDKKAPAA--- 82
Db 51 SSGSAASSGSGSASTTNSIEAKVNSDLVRNGRISGAGVYKVIHHRPSRLVYLKVI 110
Qy 83 ---HEKIQ---TYEQLVLDHENVIKLHKYMLDTPSEACARVIFITYVSSGSIKQPL 135
Db 111 YGNHEETVRQICREIEILRDVNHVNVKHEMFQNGE---IQVLFEFMKGSLE--- 163
Qy 136 KTKKKNHKAMNANAKR-----WCTQILASFLHACSPPIHGNLTSTPIFIQNGGLK 190
Db 164 -----GAHVMEQOLADISROIISGLAYLH--SHIVIRDIKPSNLINSKAVK 211
Qy 191 IGSVW--HRIFSNALRPPALPDULRSPIAREEELNHLHFPPE-----YGEVADGTA 242
Db 212 IADFGVSRILAQMDPC-----NESVGTIAYMSPERINTDLOCKY--DGYA 256
Qy 243 VDFISFGWCALEMAV---LEIQTNGDTRVTEBALARAR---HSLSDPMKEFFILCCL 294
Db 257 GDWISIGVILBEFLGRFPFPVSRQGDWASIMCAICMSQPEAPATVASEPFRHFISSCIQ 316
```

```
Qy 295 RDPARRPSAHSLLFHRVLFVYHSLKTLAAHCFIOHYLMPEVNVBEKTKAMD LHA VLAEL 354
Db 317 REFGKRRSAMQUL-----QHPFILRASPSQNR--PQNLHQULPP-- 354
Qy 355 PRP 357
Db 355 PRP 357

RESULT 3
US-11-134-563-16
; Sequence 16, Application US/11134563
; Publication No. US20050287569A1
; GENERAL INFORMATION:
; APPLICANT: Leong, John M.
; APPLICANT: Campellone, Kenneth G.
; TITLE OF INVENTION: ESPFNU NUCLEIC ACIDS AND PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 07917-280001
; CURRENT APPLICATION NUMBER: US/11/134,563
; CURRENT FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: US 60/573,600
; PRIOR FILING DATE: 2004-05-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-134-563-16

Query Match 7.2%; Score 191; DB 7; Length 545;
Best Local Similarity 24.6%; Pred. No. 3.2e-09;
Matches 84; Conservative 60; Mismatches 115; Indels 82; Gaps 16;

Qy 4 PEPARPRAREREREREDESEDSLTLE-----SPCGWQKREOVNOGNMPLQST 55
Db 221 PPDALTTRTEKQKKPKRKSDEE--ILELRISIVSGDPKKKY--TRFKIGGQ---ASGTV 284
Qy 56 FLANDTEBGEVYVMNELHFGDKKAPAAHEBKIQVFEQLVLY--DHNIYKLAHKYMLDT 112
Db 285 YTAMDVATGOEVAIKQNMVL-----OOQPKKELIINELVWRKNKNINIVA---YIDS 333
Qy 113 SEACARVIFITEYVSSGSLKQFLKTKKKNHKAMNANAKRMTQILSALSFHACSPPI 172
Db 334 YLVGDELWVMEYILAGSLTDTVYET-----CMDEGQIAVCRBCLQALEFLH--SNQVT 386
Qy 173 HGNLTSDPTFIQHNGLIKIGSVWHRIFSNALRPPALPDULRSPIAREEELNHLHFP-- 230
Db 387 HRDISDNIILGMDSVLTPFG---FCAQI-----TPQSKRSTVGTVPYMYA 432
Qy 221 PPEYGEVADGTAVIDIFSFGWCALEM-----AVLEIQTNGDTRVTEBALARA 276
Db 433 PEVVTAKAYGPKVDIWSIGIWAIEIMEGEPPYLNENPLRALYLLIATNG---TPE----- 483
Qy 277 RHLSDPM-----WSEFILCCLADPARRPSAHSLLFHRV 312
Db 484 ---LQNPFKLSAIFRDFILNRCLMDVERKRSKELLQHOFL 521
```

```
RESULT 4
US-10-661-426-4
; Sequence 4, Application US/10661426
; Publication No. US20050262584A1
; GENERAL INFORMATION:
; APPLICANT: Sneen, Jen
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Asai, Tsuneaki
; APPLICANT: Tena, Guillaume
; TITLE OF INVENTION: Master Activators of Pathogen Responsive
; TITLE OF INVENTION: Genes
; FILE REFERENCE: 00786/397003
```



```
Db 206 VYLLIEYAPLGTAVRELQKSKFDEQRYA-----TYITELANALSYCH--SKRYIHRDIKP 259
Qy 179 DTIFIONGLIKIGSVWHRISNLRPTALPDRLRSIRAREBELNLHFFPEY--GE 236
Db 260 ENLLGSAGELKID-----FGWSVHAPSSRRTL-----CGTLDYLPPEMIEGR 304
Qy 237 VADGTAVDIFSGMCALEMAVALLEIQNGDT--RVTEEAIAARASLSD---PNRBEFTLCC 292
Db 305 MHD-EKVDLWSLGLCYEFLVKGKPPFEANTYQETTKRISRVEFTFPDFTGARDLSRL 363
Qy 293 LARDPARPSASHLFF 309
Db 364 LKHNPSQRPMLREVLEH 380

RESULT 10
US-10-770-726-83
; Sequence 83, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: AM101079 (031896-010000)
; CURRENT APPLICATION NUMBER: US/10770.726
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-726-83

Query Match 6.3%; Score 168.5; DB 6; Length 403;
Best Local Similarity 24.0%; Pred. No. 2e-07;
Matches 76; Conservative 61; Mismatches 111; Indels 49; Gaps 14;

Qy 2 AAEPAARRARERE---REREDESEDSLDLESPCGRMQRRREOVNPGLOSTFLA 58
Db 104 SAEENNPD---EEELASKQNEESKKRQWALEDEFEIGR-----PLGKGR--FGNYVLA 150
Qy 59 MDTEBGEVYVWNEIHFGRKAFAPAAHEEKIQTFEQLVLDHPNIVKLHKYMLDTSACAR 118
Db 151 REKQSKRTLLAKVL--FPAOLEKAGVEHQLRREVEIQSHLHPNILRLYGFHDAT---R 205
Qy 119 VITTEYVSSGSLKQPIKTKKKNHKNAMAMKRWCTQIISLSFLHACSPPIIHGNLTS 178
Db 206 VYLLIEYAPLGTAVRELQKSKFDEQRYA-----TYITELANALSYCH--SKRYIHRDIKP 259
Qy 179 DTIFIONGLIKIGSVWHRISNLRPTALPDRLRSIRAREBELNLHFFPEY--GE 236
Db 260 ENLLGSAGELKID-----FGWSVHAPSSRRTL-----CGTLDYLPPEMIEGR 304
Qy 237 VADGTAVDIFSGMCALEMAVALLEIQNGDT--RVTEEAIAARASLSD---PNRBEFTLCC 292
Db 305 MHD-EKVDLWSLGLCYEFLVKGKPPFEANTYQETTKRISRVEFTFPDFTGARDLSRL 363
Qy 293 LARDPARPSASHLFF 309
Db 364 LKHNPSQRPMLREVLEH 380

RESULT 11
US-10-770-726-85
; Sequence 85, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
```

```
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATIN
; FILE REFERENCE: AM101079 (031896-010000)
; CURRENT APPLICATION NUMBER: US/10770.726
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 85
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-726-85

Query Match 6.3%; Score 168.5; DB 6; Length 403;
Best Local Similarity 24.0%; Pred. No. 2e-07;
Matches 76; Conservative 61; Mismatches 111; Indels 49; Gaps 14;

Qy 2 AAEPAARRARERE---REREDESEDSLDLESPCGRMQRRREOVNPGLOSTFLA 58
Db 104 SAEENNPD---EEELASKQNEESKKRQWALEDEFEIGR-----PLGKGR--FGNYVLA 150
Qy 59 MDTEBGEVYVWNEIHFGRKAFAPAAHEEKIQTFEQLVLDHPNIVKLHKYMLDTSACAR 118
Db 151 REKQSKRTLLAKVL--FPAOLEKAGVEHQLRREVEIQSHLHPNILRLYGFHDAT---R 205
Qy 119 VITTEYVSSGSLKQPIKTKKKNHKNAMAMKRWCTQIISLSFLHACSPPIIHGNLTS 178
Db 206 VYLLIEYAPLGTAVRELQKSKFDEQRYA-----TYITELANALSYCH--SKRYIHRDIKP 259
Qy 179 DTIFIONGLIKIGSVWHRISNLRPTALPDRLRSIRAREBELNLHFFPEY--GE 236
Db 260 ENLLGSAGELKID-----FGWSVHAPSSRRTL-----CGTLDYLPPEMIEGR 304
Qy 237 VADGTAVDIFSGMCALEMAVALLEIQNGDT--RVTEEAIAARASLSD---PNRBEFTLCC 292
Db 305 MHD-EKVDLWSLGLCYEFLVKGKPPFEANTYQETTKRISRVEFTFPDFTGARDLSRL 363
Qy 293 LARDPARPSASHLFF 309
Db 364 LKHNPSQRPMLREVLEH 380

RESULT 12
US-11-109-156-12
; Sequence 12, Application US/1109156
; Publication No. US20050250144A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isegai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Wakamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-ichi Punahashi
; APPLICANT: Chiaki Senoo
; APPLICANT: Jun-ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/11/109,156
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: US/10/060,065
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 14, 2006, 00:46:17 ; Search time 385 Seconds
(without alignments)
1065.817 Million cell updates/sec

Title: US-10-618-941-99
Perfect score: 2670
Sequence: 1 MAAPAPAPRARERERERED.....RMKLAFLSTLTKYKGTGA 507

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 6038814 seqs, 404674181 residues
Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=xlp
-q=/cg2_1/USPTO.spool_p/US10618941/runac.12012006.070207.14416/app_query.fasta_1.647
-DB=Published Applications NA New -OPMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=apc -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=local -OUTFMT=prc -NOR=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USR=US10618941 @CGN_1_1_184 @runac.12012006.070207.14416
-NCPUG=1 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New.*

1: /cg2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
2: /cg2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cg2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cg2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
5: /cg2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
6: /cg2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
7: /cg2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /cg2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
9: /cg2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
10: /cg2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	413	15.5	540	7	US-11-128-061-2970 Sequence 2970, Ap
2	413	15.5	540	7	US-11-128-061-6612 Sequence 6612, Ap
3	383.5	13.5	669	6	US-10-623-155-23 Sequence 23, Appl
4	359.5	13.5	6812	6	US-10-953-054-44 Sequence 44, Appl
5	302.5	11.3	7555	7	US-11-136-527-688 Sequence 688, Appl
6	199	7.5	2838	6	US-10-770-726-27 Sequence 27, Appl
7	199	7.5	4529	6	US-10-821-234-285 Sequence 285, Appl
8	192	7.2	1101	6	US-10-661-426-6 Sequence 6, Appl

9	191	7.2	1400	7	US-11-136-527-6476 Sequence 6476, Ap
10	191	7.2	1638	7	US-11-134-563-15 Sequence 15, Appl
11	191	7.2	2539	7	US-11-136-527-2380 Sequence 2380, Ap
12	190	7.1	1101	6	US-10-661-426-3 Sequence 3, Appl1
13	185.5	6.9	1964	7	US-11-136-527-3130 Sequence 3130, Ap
14	180	6.7	2241	7	US-11-136-527-2475 Sequence 2475, Ap
15	173	6.5	2253	6	US-10-770-726-39 Sequence 39, Appl
16	173	6.5	2253	7	US-11-000-688-124 Sequence 124, Ap
17	173	6.5	2347	6	US-10-770-726-41 Sequence 41, Appl
18	172.5	6.5	2033	6	US-10-523-477-4 Sequence 4, Appl1
19	170.5	6.4	1047	6	US-10-661-426-11 Sequence 11, Appl1
20	167.5	6.3	1047	6	US-10-661-426-8 Sequence 8, Appl1
21	166	6.2	2605	7	US-11-000-688-1480 Sequence 1480, Ap
22	157	5.9	3911	9	US-11-004-057-5 Sequence 5, Appl1
23	154	5.8	3260	9	US-11-004-057-1 Sequence 1, Appl1
24	154	5.8	5253	9	US-11-004-057-3 Sequence 3, Appl1
25	150	5.6	3845	7	US-11-136-527-2764 Sequence 2764, Ap
26	148	5.5	4453	7	US-11-000-688-1248 Sequence 1248, Ap
27	147.5	5.5	3293	6	US-10-770-726-43 Sequence 43, Appl
28	145	5.4	2248	6	US-10-485-517-84 Sequence 84, Appl
29	145	5.4	3258	7	US-11-136-527-3018 Sequence 3018, Ap
30	144.5	5.4	1891	6	US-10-860-501-2 Sequence 2, Appl1
31	144.5	5.4	2056	6	US-10-860-501-1 Sequence 1, Appl1
32	143.5	5.4	4554	7	US-11-136-527-2255 Sequence 255, Appl
33	142.5	5.3	3837	7	US-11-136-527-90 Sequence 90, Appl
34	142	5.3	3583	6	US-10-770-726-22 Sequence 22, Appl
35	140.5	5.3	1327	6	US-10-523-477-3 Sequence 3, Appl1
36	140.5	5.3	3331	6	US-10-770-726-42 Sequence 42, Appl
37	140	5.2	3676	7	US-11-136-527-3587 Sequence 3587, Ap
38	139.5	5.2	2932	6	US-10-860-501-8 Sequence 8, Appl1
39	138.5	5.2	1235	6	US-10-770-726-4 Sequence 4, Appl1
40	138	5.2	2119	6	US-10-770-726-26 Sequence 26, Appl
41	138	5.2	2328	6	US-10-770-726-7 Sequence 7, Appl1
42	138	5.2	2328	7	US-11-000-688-154 Sequence 154, App
43	138	5.2	5464	7	US-11-136-527-3219 Sequence 3219, Ap
44	137.5	5.1	2572	7	US-11-136-527-2589 Sequence 2589, Ap
45	137.5	5.1	3527	6	US-10-770-726-35 Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-11-128-061-2970
; Sequence 2970, Application US/11128061
; Publication No. US2006003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Ham, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US/11/128,061
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2970
; LENGTH: 540
; TYPE: DNA
; ORGANISM: *Cricetulus griseus*
US-11-128-061-2970
Alignment Scores:
Pred. No.: 8,226-37 Length: 540
Score: 413.00 Matches: 74
Percent Similarity: 81.82% Conservative: 16

Best Local Similarity:	67.27%	Mismatches:	20
Query Match:	15.47%	Indels:	0
DB:	7	Gaps:	0
US-10-618-941-99 (1-507) x US-11-128-061-2970 (1-540)			
QY	17	GIUARGIUAAPGIUSERGIUAAPGIUSERAAPP1LEUENGLUGIUSERPROCYSGIYARG	36
DB	211	GAGAGGAAAGAAAGTGAAGTGTGAGATCTTGAMAGAGTCCCTGTGACGC	270
QY	37	TRPGINLYARGARGIUGIUNVALANGINGLYASMEPROGILEUGINSETRHPHE	56
DB	271	TGGCAGAAAGGAGAAAGGGAATCAGCGTAATCTACCAAGTATTGACAGTAC	330
QY	57	LEUALAMEAAPPTRHGIUGIUGLYVALGIUVALVATTPAANGIUEWHISPHEGIASP	76
DB	331	CTGGCCATGATACAGAGAAAGGTGTGAGATTGTGTGAAATAGGACAGTTCTCGAA	390
QY	77	ARGYVALPHEALAHESGIUGIULYBILGINTHYVALPHEGIUGINLEUVALLEU	96
DB	391	CCCAAGAACTACCAACTGACGAGGAAAGGTCCGCACTGTTGTAATCTCATYCAA	450
QY	97	VALAEPHISPROANILEVALYLSLEWHISLYETITPLEUAPTHUSERGIUALCYE	116
DB	451	TGGAAACATCTTACATCTGTAAGTTTCACAAATATTGGGCTAATGTTAAAGAAATAA	510
QY	117	ALARGVALIIEPHEIETHRGIUTRYVAL	126
DB	511	GCCTGGGTATTTTCATCACAGAAATCATG	540
RESULT 2			
US-11-128-061-6612			
; Sequence 6612, Application US/11128061			
; Publication No. US20060003958A1			
GENERAL INFORMATION:			
; APPLICANT: Melville, Mark W.			
; APPLICANT: Charlebois, Timothy S.			
; APPLICANT: Mounts, William M.			
; APPLICANT: Hann, Louane E.			
; APPLICANT: Sinacore, Martin S.			
; APPLICANT: Leonard, Mark W.			
; APPLICANT: Brown, Eugene L.			
; APPLICANT: Miller, Christopher P.			
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS			
; FILE REFERENCE: 01997.027701			
; CURRENT APPLICATION NUMBER: US/11/128, 061			
; PRIOR FILING DATE: 2005-05-11			
; NUMBER OF SEQ ID NOS: 7285			
; SOFTWARE: PatentIn version 3.3			
; SEQ ID NO 6612			
; LENGTH: 540			
; TYPE: DNA			
; ORGANISM: Cricetus griseus			
US-11-128-061-6612			
Alignment Scores:			
Pred. No.: 8.22e-37 Length: 540			
Score: 413.00 Matches: 74			
Percent Similarity: 61.82% Conservative: 16			
Best Local Similarity: 87.27% Mismatches: 20			
Query Match: 15.47% Indels: 0			
DB: 7 Gaps: 0			
US-10-618-941-99 (1-507) x US-11-128-061-6612 (1-540)			
QY	17	GIUARGIUAAPGIUSERGIUAAPGIUSERAAPP1LEUENGLUGIUSERPROCYSGIYARG	36
DB	211	GAGAGGAAAGAAAGTGAAGTGTGAGATCTTGAMAGAGTCCCTGTGACGC	270
QY	37	TRPGINLYARGARGIUGIUNVALANGINGLYASMEPROGILEUGINSETRHPHE	56

```

Db      271 |GGCGAAGAGGAGGAAGGTGATCATCGCCTAAATGTAACAGTATTTCACATGTGCATAC|330
Qy      57 |LeuAlaMetAspThrGluGlnGluValAlaIleTrpAsnGluIleuHisPheGlyAsp|76
Db     331 |CTGGCGCATGAGTATACAGAGGAAAGTGTCGATGTGTGTGTGAATACAGGTACTCTCGAA|390
Qy      77 |ArgIlyAlaPheAlaAlaHisGluGlnGlyValIleGlnThrValPheGluGlnIleuValLeu|96
Db     391 |CGCAAGAACACTAACAACTGCAGAGGAAAAAGCTCCGTCGAGTGTATTATAATCATTCATCAA|450
Qy      97 |ValAspHisProAsnIleValIleuHisIleuValSerTyrTrpLeuAspThrSergIleuAlaCys|116
Db     451 |TTGGAACTATCTTAACATGCTTAAGTTTTCACAAATATTGGGCTGATGTATAAGAAATTAAA|510
Qy     117 |AlaArgValIlePheIleThrGlnTyrVal|126
Db     511 |GCCCGGTGATTTTCATCATCAGAAATACAG|540

RESULT 3
US-10-623-155-23
; Sequence 23, Application US/10623155
; Publication No. US20050261166A1
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Pechham, David W.
APPLICANT: Reiter, Marc W.
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.455C20
CURRENT FILING DATE: 2003-07-17
NUMBER OF SEQ ID NOS: 560
SOFTWARE: FaSTSeq for Windows Version 4.0
SEQ ID NO: 23
LENGTH: 669
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 642..661
OTHER INFORMATION: n = A,T,C or G
US-10-623-155-23

Alignment Scores:
Pred. No.:          2,686-33           Length:        669
Score:              383.50            Matches:         83
Percent Similarity: 66.24%             Conservative:    21
Best Local Similarity: 52.87%           Mismatches:      2
Query Match:       14.36%               Indels:          13
DB:                6                   Gaps:            3

US-10-618-941-99 (1-507) x US-10-623-155-23 (1-669)
Qy     349 |AlaValLeuAlaGluIleuProArg---ProArgArgProProLeuGlnIleTrpArgTyrSer|367
Db      7 |GCCGATCTGGCTGGAATCCCTGCAGAGCACGAGAAAGAAACAGTTCAAGACTTGTACTCT|66
Qy     368 |GluValSerPheMetGluLeuAspLysPheLeuGluAspValArgAsnGlyIleTyrPro|387
Db     67 |CAGTACCACGCTCTCGAAATTAAATTCTTGGAAAGATGTCAGAAATGGAGACTATCTCT|126
Qy     388 |LeuMetAsnPheAlaAlaThrArgProLeuGlyLeuProArgValLeuAlaProProPro|407
Db     127 |CTGACAGCCTTT-----GGGCTGCTCTGGGCCCAAGACCCACAGCAG|168
Qy     408 |GluGlnVal-----GlnIlyAlaIleThrProThrProGlnProPhe|421
Db     169 |GAGGAGGTGACATCACTACCTGTGCTGCCCCCTCTGTCAACATCTCCGACACTGAAACCACT|228
Qy     422 |ApSergIleuThrArgLysValIleGlnMetGlnCysAsnLeuGlnArgSergIleuAspLys|441

```

Db 229 GAGGTGAGAGCTGCAAGGTGTGTGCTGATGCTAGTCAACATTTGAGTCCGTGAGAGGCA 288
Qy 442 AAlaArgTPrH1sLeuThrLeuLeuValLeuGluAparGluLeuH1sAArgGlnLeuThr 461
Db 289 GTCAAAACACCACTGACACTTGTGTGCTGAGTGAAGTGAACCACTGAACCGGCACTGAGC 348
Qy 462 TTAAPLeuThrLeuProThrAparSerAlaGlnAparLeuAlaSerGluLeuValH1sArgGly 481
Db 349 TGTGACCTGATGCTCAATGAGATATCCCGAGTTGGCGGCTGAGCTGTGTGCTGAGCTGGGC 408
Qy 482 PheLeuH1sGluAparAparGluLeuAlaAparLeuGlnSerThr 498
Db 409 TTCTATTAGTGAAGCTGACCAAGCGGCTGACTTCTCTGCTGAGAGAGACT 459
RESULT 4
US-10-955-054A-44
Sequence 44, Application US/10955054A
Publication No. US20050266420A1
GENERAL INFORMATION:
APPLICANT: PUSZTAI, LAJOS
APPLICANT: SYMMANS, W. FRASER
APPLICANT: HESS, KENNETH R.
APPLICANT: AYERS, MARK
APPLICANT: STEC, JAMES
TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY
FILE REFERENCE: UTXC:880US
CURRENT APPLICATION NUMBER: US/10/955, 054A
CURRENT FILING DATE: 2004-09-30
NUMBER OF SEQ ID NOS: 195
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 44
LENGTH: 6812
TYPE: DNA
ORGANISM: Homo sapiens
US-10-955-054A-44
Alignment Scores:
Pred. No.: 5 81e-29 Length: 6812
Score: 359.50 Matches: 121
Percent Similarity: 38.16% Conservative: 74
Best Local Similarity: 23.68% Mismatches: 161
Query Match: 13.46% Indels: 155
Gaps: 13
US-10-618-941-99 (1-507) x US-10-955-054A-44 (1-6812)
Qy 5 GIUProAlaProArgArgAla-----ArgGlu 13
Db 327 GAGCGCGCGCGCGCGAGAGTGGSCAGCGCGCGCGAGCGCCCAAGAGCCACAGAGAGAA 386
Qy 14 ArgGluArgGluArgGluArgGluSerGluAparGluSerThrPro 33
Db 387 CGAGCGCAGCAGCAGATGATATCGAAAGCTGAGACCAAGCGCGGAAATGTCTAAC 446
Qy 34 CysGlyArgTPrGlnArgArgGluGlnValAparGlnGlyAparMetProGluLeuGln 53
Db 447 GATGGCGGCTTTCTCAAGTTTGCATCGAAATCGCGAGAGGCTCC-----TTTAAAG 497
Qy 54 SerThrPheLeuAlaMetAparThrGluGluGluValGluValAlaValTPrAparGluLeuH1s 73
Db 498 ACGGCTCAAAAGTGTGACACTGAAACCAACCGTGAAGTCCCGCTGAGTAACTGCGAG 557
Qy 74 PheGlyAparArgGluAparAparAlaValAparGluLeuValLeuGlnThrValPheGluGln 93
Db 558 -----GATCGAAATTTAAACAAAGTGTGAGAGCGCAAGATTTAAAGAAAGAGCTGAAGT 611
Qy 94 LeuValLeuValAparThrAparSerAlaGlnValLeuH1sLeuValTPrAparThrSer 113
Db 612 TTTAAAGCTTTCAGCATCCCAATATTTGAATTTTATGATTTCCGTGAAATCCACAGTA 671
Qy 114 GlnAlaCysAlaArgValLeuPheLeuThrGluTyrValSerSerGlySerLeuValGln 133
Db 672 AAAGGAAAGAGTGCATTGTTTGTGATGAACTTATGACGCTCTGAAACACTTAAAGG 731

Qy 134 PheLeuValLeuThrValLeuValSerAlaValAparAlaArgAlaTPrValArgTPr 153
Db 732 TATCTGAAAGG-----TTTAAAGTATGATAGATCAAAAGTTCTTAAAGACTGG 779
Qy 154 CysThrGlnLeuSerAlaLeuSerPheLeuH1sAlaCysSerProProLeuLeuH1s 173
Db 780 TCCCGTCAAGTCTTAAAGTCTTCAAGTCTTCTTCAATCTCAAGTCACTCACTTCAATTCAC 839
Qy 174 GlyAparLeuThrSerAparThrLeuPheLeuGln-----HisAparGluLeuValGly 192
Db 840 CGGATCTTAAATGATGACAAACATCTTATCAACCGGCGCTGACTGCTCAGTCAAGATTGCA 899
Qy 193 SerValTPrH1sAparGluPheSerAlaLeuValArgProProThrAlaLeuProApar 212
Db 900 GACCTC-----GGT 908
Qy 213 LeuArgSerProLeuArgAlaGlu-----ArgGluGluLeuAparGluLeuH1sPhePhe 230
Db 909 CTGGCAACCTGAAAGCGGCTTCTTGGCCAAAGTGTATAGTACCCCAAGTTTCATG 968
Qy 231 ProGluTyrGluValAlaAparGluThrAlaValAparLeuPheSerPheGlyMet 250
Db 969 GCGGCTGAGATGATGAGAGAAATATGATGAATCCGTTGACGTTATGCTTTGGGATG 1028
Qy 251 CysAlaLeuGluMetAlaValLeuGluLeuGlnThrAparGly----- 264
Db 1029 TGCATGCTTGAAGTGGCTCATCTGATATCTTACTACGAGTCCCAAAATGCTGGCGAG 1088
Qy 265 ---AparThrArgValThrGluGluAlaLeuAlaArgAlaAparHisSerLeuSerAparPro 283
Db 1089 ATCTACCGTGGCTGACCACTGGGGTGAAGCCAGCCAGTTTGAACAAAGTGAATTCCT 1148
Qy 284 AsnMetArgGluPheLeuLeuCysValLeuAlaArgAparProAlaArgProSerAla 303
Db 1149 GAAGTGAAGAAATATTGAAAGATGCATACGACAAACAAAGATGAAAG----- 1199
Qy 304 HisSerLeuLeuPheH1sArgValLeuPheGluValHisSerLeuValLeuAlaVal 323
Db 1200 -----TATTCATCAAAAGCCTTTTGAC 1223
Qy 324 HisCysPheLeuGlnH1sGlnTyrLeuMetProGluAparValValGluGluValThrVal 343
Db 1224 CATGCCCTTTCCAAAGAGAAACAGAGTACGGGTGAATTAAGCAAAAGATGAT--- 1280
Qy 344 AlaMetAparLeuH1sAlaValLeuAlaGluLeuProArgProArgProProLeuGln 363
Db 1280 ----- 1280
Qy 364 TPrArgTyrSerGluValAparPheMetGluLeuAparLysPheLeuGluAparValArg--- 382
Db 1281 -----CGAAGAAATAATACCCATTAATAATATGAGCTACGATTAAGAAATTTAAGAA 1331
Qy 383 ---AparGlyLeuTyrProLeuMetAparPheAlaAlaThrArgProLeuGluLeuProArg 401
Db 1332 TTTAAAGGAAATAC----- 1346
Qy 402 ValLeuAlaProProProGluGluValGlnValAlaValThrProThrProGluProPhe 421
Db 1346 ----- 1346
Qy 422 AapSerGluThrArgLysValLeuGlnMetGlnCysValLeuGluArgSerGluAparLys 441
Db 1347 -----AAAGATTAAGACCTATTGAGTTTCTTTGATTTTGAAGAGAG----- 1388
Qy 442 AlaArgTPrH1sLeuThrLeuLeuValLeuGluAparArgLeuH1sAparGlnLeuThr 461
Db 1388 ----- 1388
Qy 462 TTAAPLeuThrLeuProThrAparSerAlaGlnAparLeuAlaSerGluLeuValH1sArgGly 481
Db 1389 -----GATGCCAAGAGATGTGCAACAAAGATGTAAAGTCTGGG 1430

```
Qy      482 PheLeuHieGluAAspAAspArgMetLysLeuAla 492
      : : : : : | | | | | : : : : :
Db      1431 TATGTCGTGAAGGTCATCAAGACCAATGGCT 1463

RESULT 5
US-11-136-527-688
; Sequence 688, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIOR APPLICATION NUMBER: US 2005-05-25
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 688
; LENGTH: 7555
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-688

Alignment Scores:
Pred. No.:      2,17e-22      Length:      7555
Score:          302.50       Matches:      115
Percent Similarity: 36.26%   Conservatve:  71
Best Local Similarity: 22.42% Mismatch:      170
Query Match:    11.33%      Indels:       157
                        Gaps:        14

US-10-618-941-99 (1-507) x US-11-136-527-688 (1-7555)

Qy      5 GluProAlaProAArgArg-----AlaArgGluArgGlu 15
      : | | | | | | | | | | | | | | | | | | | | | |
Db      873 GAGCCGCCGCCCTCCAGAGAGTGAGAGTCGACGCGTCCAGTCCCAAGACCCGCGAG 932

Qy      16 ArgGluArgGluAspGluSerGluAsp-----GluSerAspIleLeuGluGlu 31
      : | | | | | | | | | | | | | | | | | | | | | |
Db      933 GAGGAAGGACCAACCAACAGATGATCGAAGAGCTGAGACGACGACGACGAGGAAATG 992

Qy      32 SerProCysGlyArgTyrGlnLysArgGluGluGlnValAlaGlnGluLysAsnMetProGly 51
      : | | | | | | | | | | | | | | | | | | | | | |
Db      993 TCCAAATGACGCTGCTTCTTCAATTTGACATCGAAATGCGACGAGGCTCC----- 1043

Qy      52 LeuGlnSerThrPheLeuAlaMetAspThrGluGluGluValGluValAlaTyrAsnGlu 71
      : : : : : : : : : : : : : : : : : : : : : :
Db      1044 TTTAAGACGCTGTACAAAGAGMSKMYRCCGAMWTCACSTGSAAGMSGCCSTGCTGA 1103

Qy      72 LeuHiePheGlyAspArgGlyAlaPheAlaAlaHieGluGluLysIleGlnThrValPhe 91
      : | | | | | | | | | | | | | | | | | | | | | |
Db      1104 TTGCAG-----GATCGAAAGTTTAAACAATCTGAAGGCGAGAGATTYYAAGAAAGACT 1157

Qy      92 GluGlnLeuValLeuValAspHisProAsnIleValLysLeuHieLysValTyrTyrLeuAsp 111
      : | | | | | | | | | | | | | | | | | | | | | |
Db      1158 GAAATGCTAAAGGCTGTCAACATTTWMAACATTTGTCATCTATGATTCCTGGGAATCC 1217

Qy      112 ThrSerGluAlaCysenIaArgValIlePheIleThrGlnTyrValSerSerGlySerLeu 131
      : : : : : : : : : : : : : : : : : : : : : :
Db      1218 ACAAGTAAAGGAAAAAATGCAATGTTTAAAGTGAAGTAAATGACATCTGGAACACTT 1277

Qy      132 LysGlnPheLeuLysLysLysLysValValValValValValValValValValValValVal 151
      : : : : : : : : : : : : : : : : : : : : : :
Db      1278 AAAAGCTACTTAAAAAG-----TTTAAAGATGATAAATCAAAAGTTTAAAGA 1325

Qy      152 ArgTyrCysThrGlnIleLeuSerAlaLeuSerPheLeuHieAlaCysSerProProIle 171
      : | | | | | | | | | | | | | | | | | | | | | |
Db      1326 AGCTGGTGTGCGCAAAATCTTAAAGGACTTCAATTTCTCATACAGAACTCCCGAAT 1385

Qy      172 IleHieGlyAsnLeuThrSerAspThrIlePheIleGln-----HieAsnGlyLeuLys 190
      : | | | | | : : : : | | | | | | | | | | | | |
      : : : : : : : : : : : : : : : : : : : : : :
```

```
Db      1386 ATTCACGGGATCTTAAATGTGACAACATCTTTATCACTGGTCTTACCGGCTCAGTCAAG 1445
Qy      191 IleGlySerValTyrPheIleArgIlePheSerValAlaLeuArgProProThrAlaLeuPro 210
      : | | | | | | | | | | | | | | | | | | | | | |
Db      1446 ATCGAGAGAC----- 1454

Qy      211 AspAspLeuArgSerProIleArgAlaGlu-----ArgGluGluLeuArgAsnLeuHis 228
      : | | | | | | | | | | | | | | | | | | | | | |
Db      1455 YTTGGTCTGGCAACTCTTAAAGCGGCTCTTTTGGCCAAAGTGTATAGTACCCCAAG 1514

Qy      229 PhePheProGluTyrGlyGluValAlaAspGlyThrAlaValAspIlePheSerPhe 248
      : | | | | | | | | | | | | | | | | | | | | | |
Db      1515 TTTATGGCTCTCGAATGTATGAGGAAAGTACATGATCCGTGATGTTATGCTTTT 1574

Qy      249 GlyMetCysAlaLeuGluMetAlaValLeuGluIleGlnThrAsnGly----- 264
      : | | | | | | | | | | | | | | | | | | | | | |
Db      1575 GGGATGTCAAGCTTGAAATGATGCTACGTCTGAATTCATCTACAGTCCCAAAATGCT 1634

Qy      265 -----AspThrArgValThrGluGluAlaIleAlaArgAlaArgHisSerLeuSer 281
      : | | | | | | | | | | | | | | | | | | | | | |
Db      1635 GCTCAGATCTACCGTCCAGTGCACGTGAGTGAAGCCAGCCAGTTTGGACAAAGTACGA 1694

Qy      282 AspProAsnMetArgGluPheIleLeuCysCysLeuAlaArgAspProAlaArgArgPro 301
      : | | | | | : : : : : : : : : : : : : : : : :
Db      1695 ATTCTGTAAGGAGGAAATTTATGAGAGATGTATCGACAAACAAAGATGAAGA--- 1751

Qy      302 SerAlaHisSerLeuLeuPheHisArgValLeuPheGluValHisSerLeuLysLeuLeu 321
      : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : :
Db      1752 -----TATTCATCAACAAACCTT 1769

Qy      322 AlaAlaHisCysPheIleGlnIleGlnTyrLeuMetProGluAsnValValGluGluLys 341
      : | | | | | | | | | | | | | | | | | | | | | |
Db      1770 TTGAACCAATGCCCTTTTCCAGAGGAAACAGGGGTACGGGTGAATTCAGCAAAAGAAAT 1829

Qy      342 ThrLysAlaMetAspLeuHieAlaValLeuAlaGluLeuProArgProArgArgProPro 361
      : | | | | | | | | | | | | | | | | | | | | | |
Db      1830 GAT----- 1832

Qy      362 LeuGlnTyrArgTyrSerGluValSerPheMetGluLeuAspLysPheLeuGluAspVal 381
      : | | | | | | | | | | | | | | | | | | | | | |
Db      1833 -----GGAGAAAGATAGCTATCAAAATTTATGCTACGTATTTGAAAGATATT 1877

Qy      382 Arg-----AsnGlyIleTyrProLeuMetAsnPheAlaIleThrArgProLeuGlyLeu 399
      : : : : : : : : : : : : : : : : : : : : : :
Db      1878 AAGAAATTTAAAGGMAATAC----- 1898

Qy      400 ProArgValLeuAlaProProProGluGluValGlnLysAlaLysThrProThrProGlu 419
      : | | | | | | | | | | | | | | | | | | | | | |
Db      1898 ----- 1898

Qy      420 ProPheAspSerGluThrArgLysValIleGlnMetGlnCysValnLeuGluArgSerGlu 439
      : : : : : : : : : : : : : : : : : : : : : :
Db      1899 -----AAAGCAAGAAAGCTATTGAGTTTCTTTGACCTTGAGAGAG----- 1940

Qy      440 AspLysAlaArgTyrPheLeuThrLeuLeuLeuValLeuGluAspArgLeuHisValArgGln 459
      : | | | | | | | | | | | | | | | | | | | | | |
Db      1940 ----- 1940

Qy      460 LeuThrTyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuValHis 479
      : | | | | | | | | | | | | | | | | | | | | | |
Db      1941 -----GATGTACAGAAAGATGTTTCTCAAGAAATGATGTCGAG 1976

Qy      480 TyrGlyPheLeuHieGluAspArgArgMetLysLeuAla 492
      : | | | | | : : : : : : : : : : : : : : : : :
Db      1977 TCTGGTATGTCTGTGAAGGTGATCAACAAAGCAATGGCT 2015

RESULT 6
US-10-770-726-27
; Sequence 27, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
```

```

APPLICANT: Liu, Wei
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
FILE REFERENCE: AM101079 (031896-010000)
CURRENT APPLICATION NUMBER: US/10/770,726
CURRENT FILING DATE: 2004-02-04
NUMBER OF SEQ ID NOS: 48640
SOFTWARE: PatentIn version 3.2
SEQ ID NO 27
LENGTH: 2838
TYPE: DNA
ORGANISM: Homo sapiens
US-10-770-726-27

Alignment Scores:
Pred. No.:      2,94e-11      Length:      2838
Score:          199.00       Matches:     111
Percent Similarity: 38.56%    Conservative: 76
Best Local Similarity: 22.89% Mismatch:       163
Query Matchn:   7.45%        Indels:      135
DB:             Gaps:        21

US-10-618-941-99 (1-507) x US-10-770-726-27 (1-2838)

QY      2 AAlAAProGlu---ProAlAProArGuaLAArgGLUAArgGLUAArgGLUAAb 20
Db      1045 GCGGGCCCTGTCTTCTTGAGGCCTGGGCCCTCAGACCAAGCGGAGCGCACAGGA 1104

QY      21 GlUSeRglUAeRglU-SeARpILeUgLUglUSeRProGbeLYArGrTr-GlnUYa 40
Db      1105 GTATTCCTCATAGCAAGTTCCGGGGCTG-----CCCTGACSTGTGTGGAACCCA 1152

QY      40 rGaTgTgLuGlInValaIngLnIYAmeCRProJLYLeUGInSerThrPheULameTa 60
Db      1153 GCGGACCCTGGCTC-----CTACTGGACA-----CTTCATGAG 1188

QY      60 eRrThrgLugLUgLYaIGlUValITrrAnGLUeUnH----- 73
Db      1189 AT-----TGGCAAGGCTTCCAGGACATGTCATGCGCC 1224

QY      74 -----PheGLYAer-ArGYuaLApheALaInIEGLUgLU 85
Db      1225 ACCGTGGCACTGGGGCAAGCTGTGGCTCAAGAAGATGACSTGGCGAAGCAAG 1284

QY      86 LYvIIegLnThrVALPheGLInLEuVAL-----AARIsPROAenILE 102
Db      1285 AGGCGGACGCGCTTTCACAGAGGTGTATATCAAGAGACTACAGACAGAGAAATGTG 1344

QY      103 VALYLauEunHvLyTYrTRPLueARPrThSeRGLUaLCyUALarGvaILLepHe 132
Db      1345 GTGAGCAATGTACAAACACTACTCTGTGGGGAGACAG-----CTTGGGTGTCT 1392

QY      123 ThrGIuYUrVALseRSerGLYserLEuLYbelnHEuLYvYnTHryvLYvAnHIS 142
Db      1393 ATGGAATGTTCTCGGAAGAGAGCGCTTCAAGCAATGCTGACCCACAGCAGAGTAAGAG 1452

QY      143 LYvAIameCRenaLAgrALATrPLYaRGTrPYCythrGLInIEleUsERaleUSeR 162
Db      1453 GAGCAAGATCGGGCC-----GTGTGCTTGCAGAGCTGTGACAGCCCTGTGCG 1497

QY      163 PheLUuHvAlaCYSeRProFroILAIEInIEGLYAenLEuThSeRAErPrThIEphe 182
Db      1498 GTGCTCCAGGCC-----CAGGGCGTCATCCACCGGACATCAAGAGAGCACTGATCTGTG 1551

QY      183 ILeGInIVaNGIYLelUeLYvILeGLYseRvALTrrHvARGILePeSeRAEnALA 202
Db      1552 CTGACCAATGATGGCAGGGTAGAGCTGTCAAGC-----ATTGGGATTCTGC 1596

QY      203 LeuYrPRoFroThRALeUPRoARPArLeuYrSeRProILAeGALagLUaRgLU 222
Db      1597 GCCCAGGTGACGAAGAGATGCC-----CGAAGAGAG 1629

QY      223 GLUeUARgeUnHvIeRPhE-----ProBrogLUYreJLUgLVaIALaARgLU 240

```

Db	Sequence	Accession	Length
Db	TCGCTGCTGGCAGCCCTACTGTGATGGCCCCCAAGCTCATCTCCGCTTCCCTACAGCG	1630	1689
Qy	ThralValAerPlepheserPheGlyMetCysValaleuGluMetAlaValleuGluIle	241	260
Db	CCAGAAGTAAACATCTGGTGGCTGGGAGTAATAGGTATGTAAGATG	1690	1734
Qy	GlnThrAngIyaerPThrAryValIThrGluGluAlaIleAlaArgAlaArgHisSerIeu	261	280
Db	---GTGACGAGAGAGCCCCCTACTTCAACGAGCCACCCCTCAAGGCATGATGATGTT	1735	1791
Qy	SerAerP	281	287
Db	CGGACCAACTGGCCACCCCGACTGAAGAACTGCAACAGGTGTGGCCATCTTGAAGGCG	1792	1851
Qy	PheIleIeuCysCysIleuAlaArgAerProAlaArgProSerAlaHisSerIeuIeu	288	307
Db	TTCTGTGACCGCCCTGTGCTGTGCGAGACCTCTCCACGAGGCGACGAGCCGAGCTGTG	1852	1911
Qy	PheHisArgValIeuPheGluValHisSerIeuIeuIleuAlaAlaHisCysPheIle	308	327
Db	AAGCACCATTCTGTGGCAAGGAGGAGGCGCGCTGCAAGCATGTGTGCC	1912	1959
Qy	GlnHisGlnIlyIleuMetProGluAerValValGluGluIlyThrIlyValMetAerPle	328	347
Db	-----CTCATGCGCGAGAAACCGCACAGATGAGGCGCCAGCGCCTTCC	1960	2002
Qy	uHisAlaValIleuAlaGluIeuProAerProAerAerProProIeuGlnIlyAerGlyrse	347	367
Db	-----CTCAACCAAAAGAGCCCGCGGTACACCCCGCCCCC	2003	2040
Qy	rgluVal	367	369
Db	TGAGGCGCAGTAGGGGCGCAGGCGCTCCACATCTCCAGCCCGGAGATGTCTCGCGTGGC	2041	2100
Qy	SerPheMetGluIeuAerPlePheIeuGluAerValIArgAngIlyIleTyPr	370	387
Db	ACCACTCTCTGTGTGGGAGTAAAGAACCTTA	2101	2154
Qy	oleuMetAerPheAlaIleAerPProIeuIleuIleuProAerValIleuAlaProp	387	406
Db	TTCTGTGACTTTTGAAGAAACACAGGACTCTGTGGAGCAAGGAGCTCCAGAGACCC	2155	2214
Qy	roProGluGluValGlnIlyAlaIlyThrProThrProGluProPheAerSerGluThra	406	426
Db	CACCTCTGGGACA	2215	2256
Qy	rgluys 427	426	427
Db	GGAAG 2261	2257	2261


```

FILE REFERENCE: 00786/397003
CURRENT APPLICATION NUMBER: US/10/661,426
CURRENT FILING DATE: 2003-09-12
PRIOR APPLICATION NUMBER: PCT/US 02/07650
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: US 60/275,199
PRIOR FILING DATE: 2001-03-12
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PASTESEQ for Windows Version 4.0
SEQ ID NO 6
LENGTH: 1101
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-10-618-941-99 (1-507) x US-10-661-426-6 (1-1101)

Alignment Scores:
Pred. No.: 4,11e-11 Length: 1101
Score: 192.00 Matches: 82
Percent Similarity: 37.19% Conservative: 53
Best Local Similarity: 22.59% Mismatches: 122
Query Match: 7.19% Indels: 106
DB: Gaps: 16

US-10-618-941-99 (1-507) x US-10-661-426-6 (1-1101)
QY 45 AsnGlnGlyAsnMetProGlyLeuGlnSerThrPheLeuAlaMetAspThrGlnGly 64
DB 151 TCTAGTGGATCTGCGCCGCTCTTCTGAGTGGCGCTTCCTTCAACGACATCAAGCTCC 210
QY 65 ValGlu----- 66
DB 211 ATGAAAGCGAAGAACTATTCGATTTAGTGAAGTAAACGCTATCGAAGCGAGAGCT 270
QY 67 ---ValValTrrAsnGlnLeuHisPheGlyAspArgLysAlaPheAla----- 82
DB 271 GGAAGCGTAAACAAAGTATTCACCGCTCCAGTTCCTGATATGACATTAAGTGATG 330
QY 83 -----HisGlnGlyLysIleGln-----ThrValPheGlnGlnLeuVal 95
DB 331 TACGGTAAACACAGAGAGCTGAGAGCTGACAGCTCTGTAAGAGATCGAGATTTTACGA 390
QY 96 LeuValAspHisPheProAsnIleValLysLeuHisLysTrrPheLeuAspThrSerGluAla 115
DB 391 GATGTAAATCATCTCAAAAGCTTGTGAATGTCAAGATGTTTATCAGAACGGTGAG--- 447
QY 116 CysAlaArgValIlePheIleThrGlyTrrValSerSerGlySerLeuLysGlnPheLeu 135
DB 448 -----ATCCAGGTTTGTGCTTGAATTAAGTAAAGTCTTTAGAA----- 489
QY 136 LysLysThrLysLysAsnHisLysValAlaMetAsnAlaArgAlaTrrLysArg----- 152
DB 490 -----GTTGCTCATGCTGTGGAAGCAAGCAACATTA 519
QY 153 -----TrrCysThrGlnIleLeuSerAlaLeuSerPheLeuHisAlaCysSerProPro 170
DB 520 GCGATCTATCTGCTGAGATCTTAAAGTGTGTTAGCTTATCTCAT-----ACCCTGAC 573
QY 171 IleIleHisGlyAsnLeuThrSerAspThrIlePheIleGlnHisAsnGlyLeuIleLys 190
DB 574 ATAGTTCATCGTGAATCAAAACATCAAACTTTGATAAACTCTGCTAAACGTTAAG 633
QY 191 IleGlySerValTrr---HisArgIlePheSerAsnAlaLeuArgProProThrAlaLeu 209
DB 634 ATTGCTGATTTTGGAGTGTAGTGAATCTTGCTGACAGATTAAGATCCGTG----- 684
QY 210 ProAspAspLeuArgSerProIleArgAlaGluArgGlnGluLeuArgAsnLeuHisPhe 229
DB 685 -----AATGAATCTGCTGGAACCATTCCTTAT 711
QY 230 PheProProGlu-----TrrGlyGlnValAlaAspGlyThrAla 242
DB 712 ATGAGTCTGAGAGGATTAACATGATTTGAATCAAGGAAGTAT---GATGGTTATGCT 768

```

```

QY 243 ValAspIlePheSerPheGlyMetCysValLeuGluMetAlaVal-----Leu 258
DB 769 GGAATATTTGAGAGCTTAAAGTGTAGCATTTTGAAGTTTAACTTGGGAGGTTTCTTTC 828
QY 259 GlnIleGlnThrAsnGlyAspThrArgValThrGlnGlnValIleAlaArgAlaArg--- 277
DB 829 CCTGTAGTGAACAGAGTATTTGGGCTAGCTTATGTGTCCATTTGTATGTCTACAGCT 888
QY 278 -----HisLeuSerLeuSerAspProAlaMetArgGlnPheIleLeuCysValLeuAla 294
DB 889 CCAAGAGCTTCACAGCTGCGCTGCGCGAGTTTCGACATTTTATCTGCTGTCTTCAG 948
QY 295 ArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGln 314
DB 949 AGAAGACCGGGAAGAGAGAGTGTATGCTATGCACTATG----- 987
QY 315 ValHisSerLeuLysLeuLeuAlaAlaHisCysPheIleGlnHisGlnTrrLeuMetPro 334
DB 988 -----CAGCATCCTTTCATATTAG 1008
QY 335 GluAsnValValGlnGlnLysTrrLysValAlaMetAspLeuHisAlaValLeuAlaGlnLeu 354
DB 1009 GCAAGTCCGAGCCAGAACAGTCT---CCTCAGAAATCTACATCACTTGTGCTCTCT 1062
QY 355 ProArgPro 357
DB 1063 CCTGTCCT 1071

RESULT 9
US-11-136-527-6476
; Sequence 6476, Application US/11136527
; Publication NO. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Wyeth
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6476
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-6476

Alignment Scores:
Pred. No.: 7,83e-11 Length: 1400
Score: 191.00 Matches: 87
Percent Similarity: 42.95% Conservative: 50
Best Local Similarity: 27.27% Mismatches: 104
Query Match: 7.15% Indels: 79
DB: Gaps: 14

US-10-618-941-99 (1-507) x US-11-136-527-6476 (1-1400)
QY 24 ArgGlySerAspIleLeuGlnGlnGlyArgTrrGlnLys-----Arg 40
DB 2 CTTTAAAGAAATTAACGAGCATGTGAGTGTGCGCATCTTAAGAAAGTACACAGCAG 61
QY 40 ArgGlnGlnValAsnGlnGlyAsnMetProGlyLeuGlnSerThrPheLeuAlaMetAs 60
DB 62 CTTTCAAGAAAGATGGAAGAGT-----GCTTCAAGCAAGTACAGTACACGATGGA 112
QY 60 PheGlnGlnGlyValGlnValValTrrAsnGlnLeuHisPheGlyAspArgLysAlaPhe 80
DB 113 TGTAGCCACAGGCGAGAGTGTGCGCATTAACAGATGAACCTT----- 155
QY 80 eAlaIleHisGlnGlnLysIleGlnThrValPheGlnGlnLeuValIleVal----- 97

```

Db 156 ---CAGCAGCAGCCGAGAAAGAACTCATTATTAATGAGACTCTGGTCAATGAGGAAA 211

Qy 98 -AepHisProAenIleValIyLeuHnIyLyStYrTrPrLeuAepHrSerGIuAaCybAl 117

Db 212 CAAAAACCAACATTGTGAAAC-----TATCTGGACAGATTACTCTGGGAGA 259

Qy 117 aArgValIlePheIleThGIuTyYalSerSerGIySerLeuLybAlPheLeuLybY 137

Db 260 TGAGCTATGAGTGTGTATGGAAATACCTTGCGGAGGCTCTTAAACATGTGTGACGA 319

Qy 137 eThrIyLybAenHnIyLeuAlaMetAenAlaTgAlaTrPrLybAxyTrPrCybHrGnIl 157

Db 320 AACCC-----TGATGGATGAGGCCACATAGCAGCTGTGTGCCAGAGTg 364

Qy 157 eLeuSerAlaLeuSerPheLeuHnIaIaCySerProProIleIleHnIeGIyAenLeuTh 177

Db 365 TCTAACAAGCTTGTGAGTTCCTGCAT-----TCAACCAAGTCATTTCACAGACATCA 418

Qy 177 rGerAepHrTrIlePheIleGlnHnIaenGIyLeuIleLybAlGIySerValTrpHisAr 197

Db 419 GAGTGCATATCTCTGCGGAATGAGTGCCTGTCAATTAATGA-CTTGGATTCr 477

Qy 197 gIlePheSerAenAlaLeuAxyrProHrAlaLeuProAepAepLeuAxyrSerProIl 217

Db 478 GTGCACAGATTACTCC----- 493

Qy 217 eArgAlaGIuAxyrGIuLeuAxyr-----AenLeu-HisPheAepProGIu--T 234

Db 494 -AGAGCAGAGCAAAAGAGACCATGTGGGAATCCATTTGTGATGGCAGCTGAAGTTG 552

Qy 234 yrGIyGIuValAlaAepGIyThrAlaValAsrIlePheSerPheGIyMetCyAlaLeuG 254

Db 553 TGACAACGCAAGGCTATGACCCAGCAAGTGTGATCATCTGTCTCCGTGATTAATGGCAATTG 612

Qy 254 lJmEc-----AlaValLeuGIu 266

Db 613 AATATGTTGAGGGGGAGCCCCCACTCACTCAATGAACCTTTGAGAGCTTGTACTCA 672

Qy 260 lEglnHrAenGIyAepHrHrAxyrValThrGIuGIuAlaIleAlaArgAlaArgHnIsert 280

Db 673 TTGCTATCCAAATGGG-----ACGCCAAG-----C 696

Qy 280 euSerAepProAen-----MetArgIuPheIleLeuCybCyLeuAla 295

Db 697 TTCAGAACCCAGAGAGATTGTCACTAATTTCCGGGCACTTTTGAACCCGCTGTCTGAGA 756

Qy 295 rGAsrProAlaAxyrProSerAlaHnIsertLeuPheHnAxyrValLeu 312

Db 757 TGAATGTGAAGAAGAGCTTCAGCTTAAGAGCTACTGACAGATCAATCTCTG 809

RESULT 10

US-11-134-563-15

Sequence 15, Application US//11134563

Publication No. US20050287569A1

GENERAL INFORMATION:

APPLICANT: Leong, John M.

APPLICANT: Campellone, Kenneth G.

TITLE OF INVENTION: ESPRU NUCLEIC ACIDS AND PROTEINS AND

TITLE OF INVENTION: USES THEREOF

FILE REFERENCE: 07917-280001

CURRENT APPLICATION NUMBER: US/11/134,563

CURRENT FILING DATE: 2005-05-20

PRIOR APPLICATION NUMBER: US 60/573,600

PRIOR FILING DATE: 2004-05-20

NUMBER OF SEQ ID NOS: 26

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 15

LENGTH: 1638

TYPE: DNA

ORGANISM: Homo sapiens

US-11-134-563-15

Alignment Scores:

Pred. No.:	1e-10	Length:	1638
Score:	191.00	Matches:	90
Percent Similarity:	42.57%	Conservative:	56
Best Local Similarity:	26.24%	Mismatches:	112
Query Match:	7.15%	Indels:	86
DB:	7	Gaps:	16

US-10-618-941-99 (1-507) x US-11-134-563-15 (1-1638)	
QY	4 ProgluPProAlaProAArgAlaargValaargluargrgluarargluarpgluarpgluSerGlu 23
Db	691 CCACCAAGTCTTTGGACC GGAAATAC TGAGAAGCAGAAAGAAAGAACCTTA AATCTCGAT 750
QY	24 AsgIuseRasppIIleLuglu-----SerProCyegly 35
Db	751 GAGAG-----ATCTTGAGAAATTACGAAGCATAGTGTGGCCGATCTTCAGAG 804
QY	36 ArgTrpGlnlyAsrGArsgluGlnValAsnIngInlYamMetProGlyLeuGlnSerThr 55
Db	805 AAATTT--ACACGGTTTGAGMAATTCGACAGAGT-----GCTTAGGCACCCGTG 852
QY	56 PheLuAlaMeaEapThrglugluGlyValGlulValalTrpaEngluLeuNhAspheGly 75
Db	853 TACACAGCAATGATGATGGCCACAGCAGAGAGGTGCCATTAAAGCAATGACT--- 909
QY	76 AspArglyValaPheaIalaHnsgluGlnlyellegInthrValPheGluGlnLeuVal 95
Db	910 -----CACACACACCCCMAAGACAGTCGATTATTAAATGACATCTCG 951
QY	96 LeuVal-----AspHisProAlaIlleValylLeuHlsystryrTrpLeuAspThr 112
Db	952 GTCATGAGCGGAAAACAAGAACCCCAAATTCGTGAT-----TACTTGACAGT 999
QY	113 SerGluAlaCyaaIaargValIllePheIlleThrgluTyvalSerSerglySerLeuLy 132
Db	1000 TACCTCGTGAGATGAGCTGTGGTGTATTGTGAATACTTCGTGGAGCTCTTGACA 1059
QY	133 GlnPheLeuLySyrrThryLyVaasnHlsyValaMetAmilaarglaIatrpLyarg 152
Db	1060 GATGTGTGACAGAAACT-----TGCAATGATGAGGCCAAATTGCACT 1104
QY	153 TrpCyThrGlnIlleLeuSerAlaLeuSerPheLeuHlsalacySerProProIlle 172
Db	1105 GTGTCCCTGTAGTGTCTGACAGCTCTGAGTCTTGAT-----TGCAACAGGTCA TT 1158
QY	173 HisGlyLeuLeuthrSerAspThrIllePheIllegInHlsangilyLeuIlleGly 192
Db	1159 CACAGAGCATCAAGAGTGAACAAATATCTGTGGATGATGATGGCTGTCAAGCTACT 1218
QY	193 SerValTPHIsaGlllePheSerAspAlaleuArgProProThrAlaleuProAsp 212
Db	1219 GA-CTTGGATTCTGTGCACAGATTAACCC----- 1247
QY	213 LeuArgSerProIlleArgAlaGluArgGluLeu-----ArgAsn-LeuHisPhePh 230
Db	1248 -----AAGACAGACCAACGAGACCATGTGTGAACCCCATCTGAT 1292
QY	230 eProProGlu--TyrglyGlnValAlaAspGlyThrAlaValAspIllePheSerPheG 249
Db	1293 GGCACACAGAGGTGTGACACGAAAGGCTATGGGCCCAAGTTGACATCTGTCCCTGG 1352
QY	249 yMeCyAlaLeuGluMet----- 255
Db	1353 CATATGGCCATCAAAATGATTGAAGGGAGCCTCCATPACTCAATGAACCTCTGAG 1412
QY	256 -AlaValleuGlnIlegInthrAsnglyAspThrArgValThrgluGlnAlaIleaIar 275
Db	1413 AGCCTTGACTCATTCGCCACCAATGG-----ACCCGAAA----- 1449
QY	275 gAlaArgHisSerLeuSerAspProan-----MetArgIuPheIle 290
Db	1450 -----CTTGACAGACCAAGAGCTGTACGATATCTTCGGGACCTTGAA 1496


```

Oy      1 MetAlaAlaProGluProAlaProArgArgAlaArgGluArgGluArgGluArg 20
Db      282 ATGGGAGCCCGGAGACCAGGAGCGGAGCCGGAGACGGAGCGGAGCGGAGGAG 341
Oy      21 GluSerGluAerGluSerAerIleuGluGluSerProCyGlyArgTrpGlnLysArg 40
Db      342 GAGAGCGAGGAGCGAGAGCGACATCTGGAGGAAAGCCCGTGTGGTCCGTGGCAAAAGCGA 401
Oy      41 ArgGluGlnValaGlnGlnGlyAserMetProGlyLeuGlnSerThrPheLeuAlaMetAer 60
Db      402 CGGGAGCGAGGTAAACCAAGGAAACATGCCAGGCTTCCAGAGCACTTCTCGCCATGAGAC 461
Oy      61 ThrGluGluGlyValaGluValaTrpAerGluLeuHisPheGlyAerAerGlyValaPhe 80
Db      462 ACGAGAGAGGGGGTGAAGTGGTGTGGAACAGCTCCACTTCGAGAGACAGAAAGGCTTTC 521
Oy      81 AlaAlaHisGluGluGlyIleGlnThrValaPheGluGlnLeuValaLeuValaAspHisPro 100
Db      522 GCGGCGGACGAGAGAGATCCAGACCGTGTCCAGACGCTGGTGTGGACCAACCCG 581
Oy      101 AsnIleValaGlyLeuHisLysTrpPheLeuAerThrSerGluAlaCyValaArgValaIle 120
Db      582 AACATCGTGAAGTGGACAAAGTACAGCTGGATACCTCGAGGCTCGCGCAGAGGTCTATC 641
Oy      121 PheIleThrGluTrpValaSerSerGlySerLeuLysGlnPheLeuLysValaThrLys 140
Db      642 TTCTTCACAGAGTACGTGTATCATCAGGACAGCTTCAGACCAATTCCTCAAAAGACCAAGAG 701
Oy      141 AsnHisLysValaMetAsnAlaArgAlaTrpLysArgTrpCySerThrGlnIleLeuSerAla 160
Db      702 AACCAACAGGCGCATGAACGCCCGGCGCTGAAAGCGCTGTGTGACGCAATCTGTCTGCG 761
Oy      161 LeuSerPheLeuHisAlaCySerProProIleIleHisGlyAsnLeuThrSerAerThr 180
Db      762 CTCAGCTTCCCTGACAGCTGACGCCCAATCATCCAGGAACTTCAACAGCAGCACACC 821
Oy      181 IlePheIleGlnHisAsnGlyLeuIleLysIleGlySerValaTrpHisArgIlePheSer 200
Db      822 ATCTTCATTCAGACACAGCGCTCATCAAGATCGCTCGTGTGACAGCAATCTTCTCC 881
Oy      201 AsnAlaLeuAerProProThrAlaLeuProAerAerAerAerAerAerAerAerAerAer 220
Db      882 AATGCTCGGCGCCCTCCACAGCACTTCAGATGATCTCCGAGGCCCATCCGCGCTAG 941
Oy      221 ArgGluGluLeuAerAerLeuHisPhePheProProGlyTrpGlyValaAlaAerGly 240
Db      942 CGAGGAGAACTTCGAACTGTGATCTTCCCGCAGAGTATGGAGAGTGGCCCATGG 1001
Oy      241 ThrAlaValaAerIlePheSerPheGlyMetCysAlaLeuGluMetAlaValaLeuGluIle 260
Db      1002 ACCGCTGTGACATCTTCTCTTGGATGTGTGCGCTGGAGATGGCTGTACTGGAAATC 1061
Oy      261 GlnThrAerGlyAerThrArgValaThrGluGluAlaIleAlaArgAlaArgHisSerLeu 280
Db      1062 CAGACCAATGGGAGACACCGGGGTCAAGAGAGGCCATTGGCTCGGCAAGCACTCGCTG 1121
Oy      281 SerAerProAerMetAerGluPheIleLeuCySerValaAlaArgAerProAlaArgArg 300
Db      1122 AGTGACCCCAACATGCGGAGATTCATCTTGTGCTCGGCGCGGAGACCTGCGCGCGG 1181
Oy      301 ProSerAlaHisSerLeuLeuPheHisArgValaLeuPheGluValaHisSerLeuLysLeu 320
Db      1182 CCTCTGCGCAACAGCTCTCTTCCACGCGGTCTTTCGAGGTGACATCTCGTGAAGCTC 1241
Oy      321 LeuAlaAlaHisCySerPheIleGlnHisGlnTrpLeuMetProGluAerValaGluGlu 340
Db      1242 CTGGAGGCCCACTGCTTTCATCCAGACCAAGTACCTCAAGCTCGAGAACTGTGTGAGAG 1301
Oy      341 LysThrLysAlaMetAerAerLeuHisAlaValaLeuAlaGluLeuProArgProArgArgPro 360
Db      1302 AAGACCAAGGCGCATGAGACTGACACGGGTCTTGGCGGAGACTTCCCGGCGCGCAGGCC 1361
Oy      361 ProLeuGlnTrpArgTrpSerGluValaSerPheMetGluLeuAerLysPheLeuGluAer 380

```

```

Db      1362 CCGCTGACGTGGCCGTACTCGGAAGTCTCTTCATGAGCTGAGCAAAATTCCTGGAGAT 1421
Oy      381 ValArgAerGlyIleTrpProLeuMetAerAerAerAlaIleThrArgProLeuGlyLeuPro 400
Db      1422 GTCAAGAAATGGAATCTAACCTAGTAACTTTGTCAACACTCGAACCCCTGGGCTGCC 1481
Oy      401 ArgValaLeuAlaProProProGluGluValaGlnLysAlaLysTrpThrProThrProGluPro 420
Db      1482 CGTGTGCGCCCAACCCCGAGAGAGTCCAAAGGCGCAAGACCCGACCCAGAGGCC 1541
Oy      421 PheAerSerGluTrpArgLysValaIleGlnMetGlnCysAerLeuGluArgSerGluAer 440
Db      1542 TTTCATCTGAGACCAAGAAAGTATCATCAATGACGTGCCAATCTCGAGAGAAAGAGAGAC 1601
Oy      441 LysAlaArgTrpHisLeuThrLeuLeuValaLeuGluAerAerAerAerAerAerAer 460
Db      1602 AAGCGCGCTGGCATCTCACTGCTCTGTGGTGTGAAAGACCGGCTGACCGGACAGCTG 1661
Oy      461 ThrTrpAerLeuLeuProThrAerSerAlaGlnAerLeuAlaAerGluLeuValaHisTrp 480
Db      1662 ACTTACGACCTGTCTCCCAACGAGACGAGCGCCAGGACCTCGCTGGAGCTGTGCATAT 1721
Oy      481 GlyPheLeuHisGluAerAerAerMetLysLeuAlaAlaPheLeuGluSerThrPheLeu 500
Db      1722 GCTTCTCTCCACGAGAGACGAGCGGATGAAGCTGGCCCTTCTCGAGAGACCTTCTC 1781
Oy      501 LysTrpArgGlyThrGlnAla 507
Db      1782 AAGTACCGTGGAGCCAGGCC 1802

RESULT 2
US-10-425-114-26870
; Sequence 26870, Application US/10425114
; Publication No. US2004003488BA1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26870
; LENGTH: 3837
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4656-003-G12_FLI
US-10-425-114-26870

Alignment Scores:
Pred. No.: 9.2e-303 Length: 3837
Score: 2584.00 Matches: 501
Percent Similarity: 88.05% Conservative: 0
Best Local Similarity: 88.05% Mismatches: 0
Query Match: 96.78% Indels: 68
DB: Gaps: 2

US-10-618-941-99 (1-507) x US-10-425-114-26870 (1-3837)
Oy      1 MetAlaAlaProGluProAlaProArgArgAlaArgGluArgGluArgGluArg 20
Db      231 ATGGGAGCCCGGAGACCAGGAGCGGAGCCGGAGACGGAGCGGAGCGGAGGAG 290
Oy      21 GluSerGluAerGluSerAerIleuGluGluSerProCyGlyArgTrpGlnLysArg 40
Db      291 GAGAGCGAGGAGCGAGAGCGACATCTGGAGGAAAGCCCGTGTGGTCCGTGGCAAAAGCGA 350

```

```

QY      41  ArgGUgUValAsnGInGlyAsnMetProGlyLeuGInSerThrPheLeuAlaMetAsp 60
DB      351  CGGAGCAAGTAAACCAAGGAAACATCCAGAGGCTTCAGAGCACTTCTCAGCAATGAGAC 410
QY      61  ThrGUgUGlyValGluValValTPAangIuLeuHspheGlyAspArglyAlaPhe 80
DB      411  ACGGAGAGAGGGGTAGAGGTGTGTGAAAGAGCTCCTCGGAAACAGAGAGGCTTC 470
QY      81  AlaAlaHspGluGluValIleGInThrValPheGluGluLeuValLeuValAspPro 100
DB      471  GGGGGCAAGAGAGAAAGATCCAGACCGGTGTGAGACAGCTGTGTGTGTGAGACACCCG 530
QY      101  AsnIleValIleLeuHspIleValTyrTrpLeuAspThrSerGluAlaCysAlaArgValIle 120
DB      531  AACATCGTAAATTGACAAAGTACTGTGAGTAACTCTGAGGCTTCGCGAGGCTCATC 590
QY      121  PheIleThrGluTyrValSerSerGlySerLeuValGlnPheLeuValSerThrIleVal 140
DB      591  TTCATCAGAGATAGGTGTGCATCAGGCACTCAGCAATTCCTCAAAAGACAAAGAG 650
QY      141  AsnHspIleValMetAsnAlaArg----- 148
DB      651  AACCAAGGCGCATGAAACCGGGGTATGGGAGCGGCTGGGGAGACCAAGGAGACAG 710
QY      148  ----- 148
DB      711  ACGGGGTGGGAGACCTCGGGAGCTGGAGTGTGAGGGGTGCGCGGCGGCTCGGACA 770
QY      148  ----- 148
DB      771  GGGGCTGGGCGAGATGCGGGGCGGCTCCGAGGCGGAGCCGCTCTCTCGCCACAC 830
QY      149  ----- 148
DB      831  CGAGCGAGAGTCTGTGCTGCGCGCGGCGGAGCTGGAAGCGGTGGAGCAGCATCTCG 890
QY      159  SerAlaLeuSerPheLeuHspAlaCysSerProProIleIleHspIleValMetThrSer 178
DB      891  TCTGGCTCAGTCTCTGTGACGCTGTGACGCGCCCAATCATCACAGGAACTGTGACAC 950
QY      179  AspThrIlePheIleGlnHspangIleuIleValIleValSerValTrpHisArgIle 198
DB      951  GACACCATTTCTATTCAGACAAACGCGCTCATCAAGATGGCTCGGTGGGACCGAATC 1010
QY      199  PheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIleArg 218
DB      1011  TTCTCCAAAT-----GCACTTCCAGATGATCTCGAAGCCCATTCGC 1052
QY      219  AlaGluArgGluGluLeuArgAsnLeuHspPheProProGluTyrGluValAla 238
DB      1053  GCTGAGCAGAGAACTTCGAAACCTGTCTTCTCCCGGAGATATGGAAGGTGGCG 1112
QY      239  AspGlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaValIleu 258
DB      1113  GATGGAGCCGCTGTGACATCTTCTCTTGGAGTGTGTGCGCTGAGATGCTGTACTG 1172
QY      259  GluIleGInThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArgHis 278
DB      1173  GAATCCAGACCAATGGGGACACCGGGGTCAAGAGAGGCCATTGCTGGCCGAGGAC 1232
QY      279  SerLeuSerAspProAsnMetCArgLupheIleLeuCysValLeuAlaArgAspProAla 298
DB      1233  TCGGAGATGACCCCAACATGGGAGATTTCATCTTGTGTGCTGGCCCGGAGACCTG 1292
QY      299  ArgArgProSerAlaHisSerIleuPheHisArgValIleuPheGluValHisSerIleu 318
DB      1293  CGCGGCGCTCTGTGCGCAAGCTCTTTCACCGGAGTCTTTGAGGTGACATCGCG 1352
QY      319  LysValLeuAlaAlaHisCysPheIleGlnHspGlnTyrLeuMetProGluAsnValVal 338
DB      1353  AAGCTCTGGGAGGCCCATGCTTATCCAGACCAAGTACTGATCTGAGAAATGTGGTG 1412

```

```

QY      339  GluGluValThrIleValMetAspLeuHspAlaValIleuAlaGluLeuProArgProArg 358
DB      1413  GAGGAAAGACCAAGGCAATGACCTGACAGCGGTCTTGCGAGCTTCCCGGCCCCCG 1472
QY      359  ArgProProLeuGlnTrpArgTyrSerGluValSerPheMetGluLeuAspIlePheLeu 378
DB      1473  AGGCCCCGCTCAGTGGGCGTACTCGAAGCTCTCTTCAAGAGCTGAGCAAAATTCCTG 1532
QY      379  GluAspValArgAsnGlyIleTyrProLeuMetAsnPheAlaAlaThrArgProLeuGly 398
DB      1533  GAGGATGTGAGAAATGAAATTAACCACTGAAGAACTTTGACAGCACTGAGACCTGGGG 1592
QY      399  LeuProArgValLeuAlaArgProProGluGluValGlnIleValIleValThrProThrPro 418
DB      1593  CTGCCCCGTGTCTGCGCCACCCCGAGAGAGTCAAAAGGCCAAAGCCCGGAGCGCA 1652
QY      419  GluProPheAspSerGlyThrArgIleValIleGlnMetGlnCysAsnLeuGluArgSer 438
DB      1653  GAGCCTTTGACTGTGAGACCAAGAAAGTCAATCCAGATGCAATGCACTTGAGAGAAC 1712
QY      439  GluAspIleValArgTyrPheIleuThrIleuLeuValIleuGluAspArgIleHisArg 458
DB      1713  GAGCAAGCGCGCTGGCATCTCACTGTGTGTGCTGGAAGACCGGCTGCACCG 1772
QY      459  GlnLeuThrTyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuVal 478
DB      1773  CAGCTGACTTACGACCTGTCTCCAAAGCAGCGCCGAGACCTGCGCTGAGGCTCGTG 1832
QY      479  HisTyrGlyPheLeuHspIleGluAspAspArgMetIleValAlaIlePheLeuGlnSerThr 498
DB      1833  CACTATGCTTCTCTCCACAGAGACGACCGGATGAAGTGTGCGGCTCTTCTGAGAGAC 1892
QY      499  PheLeuIleTyrArgGlyThrGlnAla 507
DB      1893  TTCTCAAGTACCGTGGGACCCAGGCC 1919

```

RESULT 3

```

US-10-425-114-26848
; Sequence 26848, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack B
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26848
; LENGTH: 3838
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4654-015-F10_FLI
US-10-425-114-26848

```

Alignment Scores:

```

Pred. No.: 9,2e-303 Length: 3838
Score: 2584.00 Matches: 501
Percent Similarity: 88.05% Conservative: 0
Best Local Similarity: 88.05% Mismatches: 0
Query Match: 96.78% Indels: 68
DB: 7 Gaps: 2

```

US-10-618-941-99 (1-507) x US-10-425-114-26848 (1-3838)

```

QY      1  MetAlaAlaProGluProAlaProArgArgAlaArgGluArgGluArgGluArgPro 20
      |||

```


US-10-618-941-99 (1-507) x US-10-425-114-26873 (1-3846)

QY 1 MetAlaAlaProGluProAlaProAlaGAlaArgGluArgGluArgGluArg 20

Db 216 ATGGGCGGCCCCGAGAGCGGCGCGGAGCGGCGCGGAGCGGAGCGGAGCG 275

QY 21 GluSerGluAspGluSerAspIleLeuGluGluSerProCysGluYArgTrpGlnYAsp 40

Db 276 GAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 335

QY 41 ArgGluGlnValAsnGlnGlnValAsnMetProGlyLeuGlnSerThrPheLeuAlaMetAsp 60

Db 336 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 395

QY 61 ThrGluGluGlyValGluValValTrpAsnGluLeuHisPheGlyAspArgIleAlaPhe 80

Db 396 ACGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 455

QY 81 AlaAlaHisGluGluValIleGlnThrValPheGluGlnLeuValLeuValAspPro 100

Db 456 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 515

QY 101 AsnIleValIleLeuHisIleValTrpLeuAspThrSerGluAlaCysAlaArgValIle 120

Db 516 AACATCGTGAAGTTCGACAAAGTACGCGCTGATACCTCTGAGGCGCTGCGGAGGAGTATC 575

QY 121 PheIleThrGluValIleSerSerGlySerLeuValGlnPheLeuValIleValThrIleVal 140

Db 576 TTCATCACAGAGTACGTTCATCAGGCGAGCTCAAGCAATTCCTCAAAAAGCAAGAGAG 635

QY 141 AsnHisIleValIleMetAspAlaArg----- 148

Db 636 AACCAAGAGGCGCATGAAAGCGCGGATGAGGAGCGGAGCGGAGCGAGCAAGAG 695

QY 148 ----- 148

Db 696 ACGGGGTTGGGGCAGCCTCGGGGAGCTGGGATGATGAGGGGGTCCCGGGGCGCTCGGACA 755

QY 148 ----- 148

Db 756 GGGGCTGGGCGAGATGCGGGGCGGGGCTCGGAGGCGGAGCGGCTCTCGGCCACAC 815

QY 149 ----- 148

Db 816 CGACCGAGGAGTCTGCGCTCGCGCGGAGGCGGCTGAGAGCGCTGAGTCCGAGTCCG 875

QY 159 SerAlaLeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAsnLeuThrSer 178

Db 876 TCTGGGCTCAGCTTCTGCAAGCGCTGCAAGCGGCGGAGGAGGAGGAGGAGGAGGAG 935

QY 179 AspThrIlePheIleGlnHisAsnGlyLeuIleValIleValIleValIleValIleVal 198

Db 936 GACACCACTTCATTCAGGACAAAGCGCTCATCAAGATCGGCTCGGTGGGACCGAATC 995

QY 199 PheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIleArg 218

Db 996 TTCCTCAAT-----GCACTTCAGAGATCTCCGAGGCCCTCCGCTCGC 1037

QY 219 AlaGluArgGluGluLeuArgAsnLeuHisPhePheProProGluIleValIleValIle 238

Db 1038 GCTGAGCGAGAGAACTTCGAAAGCTTCGAACTTCCTCCCGGAGAGTGAAGAGGAGGAG 1097

QY 239 AspGlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaValLeu 258

Db 1098 GATGGGACCGCTGTGACATCTTCCTTCCTTGGGATGTGTGGCTGAGATGGCTGTACG 1157

QY 259 GluIleGlnThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArgHis 278

Db 1158 GAATTCGAGCAACAAAGGAGGAGCGCGGCTCAGAGGAGGAGGAGGAGGAGGAGGAG 1217

QY 279 SerLeuSerAspProAspMetArgGluPheIleLeuCysValLeuAlaArgAspProAla 298

Db 1218 TCGCTGAGTGAACCCCAACATGGGGAGTTCATCTTTGCTGCTGGCCCGGAGCCCTGCC 1277

QY 299 ArgArgProSerAlaHisSerIleuPheHisArgValLeuPheGluValHisSerLeu 318

Db 1278 CCGCGGCTCTGCGCCACAGACCTCTCTTCCACCGGCTCTCTTCAGAGTGAATCTGCTG 1337

QY 319 LysLeuLeuAlaHisArgPheIleGlnHisGlnIleValLeuMetProGluAsnValVal 338

Db 1338 AAGCTCTGGAGCCCATCTGCTTATCCAGCACAGTACTCATGCTTGAAGATGTGTG 1397

QY 339 GluGluValThrLysAlaMetAspLeuHisAlaValLeuAlaGluLeuProArgProArg 358

Db 1398 GAGAGAAAGACCAAGGCGCATGACCTGACAGCGGCTTGGCGGAGCTTCCCGGCGCCGC 1457

QY 359 ArgProProLeuGlnTrpArgTrpSerGluValSerPheMetGluLeuAspIlePheLeu 378

Db 1458 ACGCCCGGCTGAGGCGGCTGATCTCGAAGGCTCTCTCATGAGCTGGAATTCCTG 1517

QY 379 GluAspValArgAsnGlyIleTrpProLeuMetAspPheAlaIleThrArgProLeuGly 398

Db 1518 GAGATGTGAGAAATGAAATCTAACCACTGATGAATCTTGGCAAGCTGAGCCCTGGAGG 1577

QY 399 LeuProArgValLeuAlaProProProGluGluValGlnLysAlaIleValThrProThrPro 418

Db 1578 CTGCCCCGTGCTCTGCCCCCAACCCCGAGAGGTCGCAAAAGCCCAAGACCCCGAGCCA 1637

QY 419 GluProPheAspSerGluThrArgValIleGlnMetGlnCysAsnLeuGluArgSer 438

Db 1638 GAGCCCTTGAATCTGAGACAGAAAGGTCATCAGATGCAATCTGAGAGAGAAC 1697

QY 439 GluAspValArgTrpHisIleuThrLeuLeuValLeuGluAspArgLeuHisArg 458

Db 1698 GAGACAAAGCGCGCTGGATCTCACTGCTTGTGTGTGAAGACCGGCTGACCG 1757

QY 459 GlnLeuThrTrpAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuVal 478

Db 1758 CAGCTGACCTAGCATCTGCTCCCAAGACGACGCGGAGGAGCTCGGCTCGAGCTCGTG 1817

QY 479 HisTrpGlyPheLeuHisGluAspAspArgMetLysLeuAlaIlePheLeuGluSerThr 498

Db 1818 CACTATGGCTTCTCCAGAGGAGCAGCGAGTGAAGTGGCGGCTTCTGAGAGAGAC 1877

QY 499 PheLeuValTrpArgGlyThrGlnAla 507

Db 1878 TTCCTCAAGTACGTGGAGCCAGGCG 1904

RESULT 5

US-10-425-114-28974

Sequence 28974, Application US/10425114

Publication No. US2004003488A1

GENERAL INFORMATION:

APPLICANT: Liu, Jindong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E.

APPLICANT: Tabaska, Jack E.

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-2153131B

CURRENT APPLICATION NUMBER: US/10/425,114

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 28974

LENGTH: 3128

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Clone ID: LIB4809-016-A10_FLI

US-10-425-114-28974

Alignment Scores: 1.65e-296 Length: 3128

Score: 2531.50 Matches: 499
Percent Similarity: 83.47% Conservative: 1
Best Local Similarity: 83.31% Mismatches: 1
Query Match: 94.81% Indels: 99
DB: 7 Gaps: 3
US-10-618-941-99 (1-507) x US-10-425-114-28974 (1-3128)

Qy 1 MetAlaAlaProGluProAlaProArgArgAlaArgGluArgGluArgGluArgGluArg 20
Db 120 ATGGGGGCCCCGGAGCCGGCCGGAGGGGGCCGGAGACGGAGACGGAGACGGAGAC 179
Qy 21 GluSerGluAerGluSerAerPileuGluGluSerProCySGlyArgTrpGlnLysArg 40
Db 180 GAGACGAGAGACGAGAGACATCTCGAGAGAAACCCGTGTGTGTGTGTGTGTGTGTGT 239
Qy 41 ArgGluGluValAerGlnGlyAerMetProGluGlnGlnSerTrpPheLeuAlaMetAer 60
Db 240 CGGGAGCAGGTAAACCAAGGAGACATGCCAGGGCTTCAGAGACCTTCTAGCCATGGAC 299
Qy 61 ThrGluGluGlyValGluValValTrpAerGluLeuMetPheGlyAerArgLysAlaPhe 80
Db 300 ACGGAGAGGGGGGTAGAAGTGTGTGTGAACGAGCTCCACTTCGAGAGACGAGAGCCCTTC 359
Qy 81 AlaAlaHisGluGluLysIleGlnThrValPheGluGlnLeuValLeuValAerPhePro 100
Db 360 GCGGCGCACGAGAGAGACATCCAGCCGTTCAGACAGCTGTGTGTGTGTGTGTGTGTGT 419
Qy 101 AenIleValLysLeuMetLysTrpLeuAerTrpSerGlnAlaCyValAerValIle 120
Db 420 AACATCGGAAGTTCACAGATACGTGCTGATACCTTCAGAGCCCTGCGAGGGGTATC 479
Qy 121 PheIleThrGluTrpValSerSerGlySerLeuLysGlnPheLeuLysTrpLys 140
Db 480 TTCATCAAGAGTACGTGTATCAAGCAGCCCTCAAGCAATTCCTCAAAAGACCAAGAG 539
Qy 141 AenHisLysAlaMetAerAlaArg----- 148
Db 540 AACCAAGAGCCATGAACGCCCGGATATGGGAGCGGGCTGGGAGCAGCCAGGGAGCAGG 599
Qy 148 ----- 148
Db 600 ACGGGGTGGGGCAGACTCGGGAGCTGGGATGTGAGGGGGTGCCTGGCGCTCGAGCA 659
Qy 148 ----- 148
Db 660 GGGGCTGGGCGAGATGCGGGGCGGGCTCGCAGGCCAGCCGCTTCCTGCGCCACCC 719
Qy 149 -----AlaTrpLysArgTrpCySerTrpGlnIleLeu 158
Db 720 CGACCGAGAGAGT 779
Qy 159 SerAlaLeuSerPheLeuMetAlaCySerProPheIleHisGlyAerLeuTrpSer 178
Db 780 TCTCGCTCAGACTTCTGTCAAGCTGTGAGCCCAATATCATCAAGGAGACCTGACGAGC 839
Qy 179 AserThrIlePheIleGlnHisAerGlyLeuLysIleGlySerValTrpHisArgIle 198
Db 840 GACACCATCTTATTCAGACACACGCGCTCATCAAGATCGGCTCGTGTGACCGAGATC 899
Qy 199 PheSerAerAlaLeuArgProProThrAlaLeuProAerAerPheLysArgProIleArg 218
Db 900 TTCTCCAAAT-----GCACTTCCAGATGATCTCCAGAGCCCATCTCGC 941
Qy 219 AlaGluArgGluGluLeuArgAerLeuMetPhePheProProGluTrpGlyValAla 238
Db 942 GCTGAGCAGAGAGAACTTCGAACTTCGAACTTCTCCCAAGAGTATGAGAGGTGGCC 1001
Qy 239 AserGlyThrAlaValAerPhePheSerPheGlyMetCyValAerGluMetAlaValLeu 258
Db 1002 GATGGAGACCGCTGTGAGACATCTTCTCTTGTGGATGTGTGTGTGTGTGTGTGTGT 1061
Qy 259 GluIleGlnThrAerGlyAerThrArgValThrGluGluAlaIleAlaArgAlaArgHis 278

Db 1062 GAATTCAGACCAATGGGGAGACCCGGGTCAAGAGAGGCCATTTGCTCGGCCAGGAC 1121
Qy 279 SerLeuSerAerProAerMetArgGluPheIleLeuCySerLeuAlaArgAerProAla 298
Db 1122 TCGCTGAGTACCCCAACATGCGGAGATTCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1181
Qy 299 ArgAerProSerAlaHisSerLeuLeuPheHisArgValLeuPheGluValHisSerLeu 318
Db 1182 CGCGGCGCTTGTGCCACAGCTCTCTTCCACGCGGTGTCTTCTGAGGTGACCTGCTG 1241
Qy 319 LysLeuLeuAlaAlaHisCySerPheIleGlnHisGlnTrpLeuMetProGluAerVal 338
Db 1242 AAGCTCTGGCAGCCCACTGCTTATCATAGACCAACAGTACCTCAAGCTGAGAAATGTGTG 1301
Qy 339 GluGluLysThrLysAlaMetAerPheLeuAlaValLeuAlaGluLeuProAerProArg 358
Db 1302 GAGGAGAAAGACCAAGGCCATGACCTGACGCGGTCTTGGGGAGCTTCCCGGCCCGGC 1361
Qy 359 ArgProProLeuGlnTrpArgTrpSerGluValSerPheMetGluAerLysPheLeu 378
Db 1362 AGGCCCCGTGCAGTGCGGTACTGCAAGTCTCTTCAATGAGCTGAGCAAAATTCCTG 1421
Qy 379 GluAerValArgAerGlyIleTrpProLeuMetAerPheAlaAlaTrpArgProLeuGly 398
Db 1422 GAGGATTCAGAGATGGAATCTACCTCATGAACTTTGACGCCACTGACCCCTGGGG 1481
Qy 399 LeuProArgValLeuAlaProProProGluGluValGlnLysValLysTrpProThrPro 418
Db 1482 CTGCCCCGTGTGTGCGCCCAACCCCGAGAGAGGTCCAAAGGCCAAGACCCCGACGCA 1541
Qy 419 GluProPheAerSerGluThrArgLysValIleGlnMetGlnCySerLeuGluAerSer 438
Db 1542 GAGCCCTTGACTCTTGAGACAGAAAGTATCATCAAGTGAAGTCAACTGAGAGAGAGC 1601
Qy 439 GluAerLysAlaArgTrpHisLeuThrLeuLeuValLeuGluAerArgLeuHisArg 458
Db 1602 GAGGACAAAGCGCGCTGTGCATCTCATCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 1661
Qy 459 GlnLeuThrTrpAerLeuLeuProThrAerSerAlaGlnAerPheLeuAlaSerGluLeuVal 478
Db 1662 CAGCTGACCTACGACTGTCTCCCAAGCAGCGCCCAAGACCTGCTGGAAGCTGTG 1721
Qy 479 HisTrpGlyPheLeuMetGlu----- 485
Db 1722 CACTATGCTTCTTCCACGA-GGTGCGTGGCGGTGCGCGCGCTGTGCGAGGGGCGC 1780
Qy 486 -----AerAerArg 488
Db 1781 GCACGGGGCAGCGCGCGCCCTCCGTCCCATGCTCCCTCTTCCGCGAGAGACCGG 1840
Qy 489 MetLysLeuAlaAlaPheLeuGluSerTrpPheLeuLysTrpArgIleThrGlnAla 507
Db 1841 ATGAAGCTGCGCGCTTCTGTGAGAGACATCTTCTCAAGTACCGTGGAGACCGAGCC 1897

RESULT 6
US-10-840-512-71
; Sequence 71, Application US/10840512
; Publication No. US20050125852A1
; GENERAL INFORMATION:
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: MANNING, GERARD
; APPLICANT: CHARIDCZAK, GLEN
; APPLICANT: GRIGORIEV, IGOR
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-1455
; CURRENT APPLICATION NUMBER: US/10/840, 512
; CURRENT FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: 60/469, 014
; PRIOR FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 239
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 71


```
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 117
; LENGTH: 3538
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-175-523-117

Alignment Scores:
Pred. No.:      1,656-276      Length:      3538
Score:          2368.00        Matches:     460
Percent Similarity: 87.17%    Conservative: 2
Best Local Similarity: 86.79%  Mismatches:  0
Query Match:    86.69%        Indels:     68
DB:             5             Gaps:       2

US-10-618-941-99 (1-507) x US-10-175-523-117 (1-3538)

QY      40  ArgArgGluGlnValAsnGlnGlyAsnMetProGlyLeuGlnSerThrPheLeuAlaMet 59
DB      1  CAGAGAGAGAAAGTAAACCAAGGAACATGCCAGGAGCTTCAGAGCACTTCCTAGCCATG 60
QY      60  AaPThrgLugLugLValGluValValTPraNGluLeuNH1aPheGlyAspArgLysAla 79
DB      61  GACACGAGAGAGGGGGTGAAGGTGTGTGGAACGAGCTTCCACTTCGAGACAGAGAGGCC 120
QY      80  PheAlaAlaNH1aGluGluValLeuGlnThrValPheGluGlnLeuValLeuValAspHis 99
DB      121  TTCGGGGGCGACGAGAGAAAGATCCAGACCGGTGTTCGAGCAGCTGTCTGGTGACAC 180
QY      100  ProAlaValLeuLeuNH1aLysPheTyrTrpLeuAspThrSerGlyAlaCysAlaArgVal 119
DB      181  CCGAACATCGTAAGTGCACAAAGTACTGGCTGTGATCTTCGAGGCTTCGCGAGGGTTC 240
QY      120  IlePheIleThrGluTyrValSerSerGlySerLeuLysGlnPheLeuLysPheThrLys 139
DB      241  ATCTTCATCAGAGATACGTGTCTGTATCAGGCACTTCAGCAATTCCTCAAAAGCCAAAG 300
QY      140  LysAsnNH1aLysAlaMetAsnAlaArg----- 148
DB      301  AAGAACCAAGAGGCATGAACGCCCGGGATATGGAGAGCGGGCTGGGAGCCACCGGGAGC 360
QY      148  ----- 148
DB      361  AGAGCGGGGTGGGAGAGCTCGGGAGCTGGAGTGTAGAGGGAGTCCCGCGGCTTCGG 420
QY      148  ----- 148
DB      421  ACAGGGGCTGGGCGAGAGATGGGGGGGCTCCGAGGCCAGCCGCTTCCTGCGGCC 480
QY      149  -----AlaTrpLysArgTyrCysThrGlnIle 157
DB      481  ACCCGACCGAGCGAGATCGTGTGCTCCGCCGCCAGCGCTGGAAGCGCTGTGTGACGAGATC 540
QY      158  LeuSerAlaLeuSerPheLeuNH1aLysSerProProIleIleHisGlyAsnLeuThr 177
DB      541  CTGTGTGGCTAGCTTCCTGTGACGCTGTGACGCCCAATCATCAAGGGAACCTTGACC 600
QY      178  SerAspThrIlePheIleGlnHisAsnGlyLeuLysIleGlySerValTPHHisArg 197
DB      601  AGCGACACATCTTATTCAGACCAACGGCTTCATCAAGATGGGCTCCGTGTGGACCGA 660
QY      198  IlePheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIle 217
DB      661  ATCTCTCCAAAT-----GCACCTTCAGATGATCTCGAAGCCCATC 702
QY      218  ArgAlaGluArgGluGluLeuArgAsnLeuNH1aSerPheProProGluTyrGlyVal 237
DB      703  CGCGGTGAGCGAGAGAACTTCGAACTTCCTTCCCGCCGAGATATGAGAGGTG 762
QY      238  AlaAspGlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaVal 257
```

```
DB      763  GCCAGTGGAGACCGGTGTGACATCTTCTTGGGATGTGTGGCGGTGAGATGCGTGA 822
QY      258  LeuGluIleGlnThrAsnGlyAspThrArgValIleThrGluGluAlaIleAlaArg 277
DB      823  CTGAAATTCAGACCAATGGGGACACCGGGTCAAGAGAGGCATTGCTCGGCGCAGG 882
QY      278  HisSerLeuSerAspProAsnMetArgGluPheIleLeuCysGlyLeuAlaArgAspPro 297
DB      883  CACTCGCTGAGTACCCCAACATCGGAGATTCTTCTTCTGCTGCGCCCGGAGCCCT 942
QY      298  AlaArgArgProSerAlaNH1aSerLeuLeuPheNH1aArgValLeuPheGluValHisSer 317
DB      943  GCCCGCGCGCCCTCTGCCCAACAGCTCTTCCACCGCGGTCTTTCGAGGTGACATCG 1002
QY      318  LeuLysLeuLeuAlaAlaNH1aLysPheIleGlnHisGlnTyrLeuMetProGluAsnVal 337
DB      1003  CTGAAGCTCTGGAGGCCACCTGCTTATCCAGACCAAGTACCTCATCCCTGAGAGATGTG 1062
QY      338  ValGluGluYsrThrLysAlaMetAspLeuNH1aValLeuAlaGluLeuProArgPro 357
DB      1063  GTGAGAGAGAAACCAAGCCATGACCTGCACCGGCTTGTGGCGAGCTTCCCGGCGCC 1122
QY      358  ArgArgProProLeuGlnTyrArgTyrSerGluValSerPheMetGluLeuAspLysPhe 377
DB      1123  CGCAGGCCCCCGCTGCAGTGGCGGTACTCGGAATCTCTTCATGAGAGCTGGAACAATTC 1182
QY      378  LeuGluAspValAlaArgAsnGlyLysTyrProLeuMetAsnPheAlaIleThrArgProLeu 397
DB      1183  CTGAGAGATGTCAAGAAATGGAATCAACCACTGATGAACCTTGGACGCCACCTGCACCCCTG 1242
QY      398  GlyLeuProArgValLeuAlaProProProGluGluValGluLysAlaLysPheProThr 417
DB      1243  GGGCTGCCCGGTGTGTGGCCACCCCGAGAGAGTCCAAAGGCCAAAGCCCGGACG 1302
QY      418  ProGluProPheAspSerGluThrArgLysValIleGlnMetGlnCysAsnLeuGluArg 437
DB      1303  CCAAGGCCCTTTGACTGTGAGACCAAGAGTCAATCCAGATGAGTGAATCTGAGAGGA 1362
QY      438  SerGluAspLysAlaArgTPHHisLeuThrLeuLeuValLeuGluAspArgLysHis 457
DB      1363  AGCCAGAGCAAGAGCGCCCTGTCATCTGCTGTGTGGTGTGCGAAGACCGGCTGCAC 1422
QY      458  ArgGlnLeuThrTyrAspLeuLeuProThrAspSerLysGlnAspLysAlaSerGluLeu 477
DB      1423  CGGAGCTGACCTAGCAGCTGTCTCCAAAGGACGCGCCAGACCTGCGCTCGAGGTTC 1482
QY      478  ValHisTyrGlyPheLeuNH1aGluAspAspArgMetLysLeuAlaAlaPheLeuGluSer 497
DB      1483  GTGCATATGTGCTTCTTCCAGAGAGCAAGCCGATGAAGCTGGCGGCTTCTGGAAGAC 1542
QY      498  ThrPheLeuLysTyrArgGlyThrGlnAla 507
DB      1543  ACCTTCTCAAGTACCGTGGGACCCAGGCC 1572

RESULT 8
US-11-099-266-117
; Sequence 117, Application US/11099266
; Publication No. US20050181433A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Paley, Michael
; APPLICANT: Rajan, Pithi
; TITLE OR INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 03235/100U795-US4
; CURRENT APPLICATION NUMBER: US/11/099,266
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: US 10/175,523
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
```

```

? PRIOR FILING DATE: 2001-06-18
? PRIOR APPLICATION NUMBER: US 60/317,828#
? PRIOR FILING DATE: 2001-09-07
? PRIOR APPLICATION NUMBER: US 60/325,150#
? PRIOR FILING DATE: 2001-09-25
? PRIOR APPLICATION NUMBER: US 60/333,047#
? PRIOR FILING DATE: 2001-11-14
? PRIOR APPLICATION NUMBER: US 60/349,936#
? PRIOR FILING DATE: 2002-01-18
? PRIOR APPLICATION NUMBER: US 60/361,834#
? PRIOR FILING DATE: 2002-03-04
? NUMBER OF SEQ ID NOS: 197
? SOFTWARE: Patencin version 3.1
? SEQ ID NO 117
? LENGTH: 3538
? TYPE: DNA
? ORGANISM: Homo sapiens
US-11-099-266-117

```

Alignment Scores:	
Pred. No.:	1,659-276
Score:	2368.00
Percent Similarity:	87.17%
Best Local Similarity:	86.79%
Query Match:	88.69%
DB:	10
	Gaps: 2
	Indels: 68
	Mismatches: 0
	Conservative: 2
	Matches: 460
	Length: 3538

US-10-618-941-99 (1-507) x US-11-099-266-117 (1-3538)

Qy	40	AtgAtggtlucInValAengInglYAemEt ProGlyLeuGInSerThrPhenLeuAlaMet	59
	1	CAAGAGAGAAAGTAAACCAAGGAAACATGCCAGGCTTCAGAGCACTTCCTAGCCATG	60
Qy	60	AaPThrGluGluGluValAlaGluValAlaTTPraengIleuHisPheGluYAspArgVala	79
Ds	61	GACACGGAGAGGGGTGAAGAGGTGTGTGGAAAGAGCTCACTTCGAGACAGGAAGGCC	120
Qy	80	PheAlaAlaHisGluGluGluValYsIIegInThrValPheGluGluInleuValLeuValAspHis	99
Ds	121	TTGGCGGCGACAGAGAGAAAGATCCAGACCGTGTTCAGACAGCTGGTGTCTGTGGACAC	180
Qy	100	ProAenIIeValIyValYeuHisIyValYtTTPLeuAaPThrSerGluAlaCysAlaArgVal	119
Ds	181	CCGAACATCTGTGAATGTGCACAAAGTACTGGCTGGATACCTCTGAGGCTTGGCCAGGCTC	240
Qy	120	IlePheIIeThrGluTyrValSerSerGlySerIleuYsGInPheLeuYsIyValThrIys	139
Ds	241	ATCTTCATCATCAGAGTACAGTCAATCAGGACGCTCAAGCAATTCCTCAAAAACACAG	300
Qy	140	LysAenHisIyValAlaMetAenAlaArg-----	148
Ds	301	AAGAACCCACAAAGGCGCATGAACCGCGGTATGGGAGCGGGCTGGGGCAGCCAGGGAGC	360
Qy	148	-----	148
Ds	361	AGGACGGGGTGTGGGGGCAAGCTCGGGGACTGGGATGCTGAAGGGGGTCCCGGCGGCTTCG	420
Qy	148	-----	148
Ds	421	ACAGGGGCTGGGCGAGGATGGGGGCGGGCTCCGAGGCCAGCGGCTCTTCGCGCC	480
Qy	149	-----AlaTrrIyValGtTrrCysThrGlnIle	157
Ds	481	ACCGGACCGAGGAGTCTGGCTCGCGCGCCAGGCGCTGAAGCGTGTGTGACGCAATC	540
Qy	158	IeuSerAlaIeuSerPheLeuHisAlaCysSerProProlIeIleHisGlyAsnIleThr	177
Ds	541	CTGCTGGCGCTGAGCTTCCTGACGCGCTGCAAGCCCCCAATCATATCAGAGGAAGCTGACC	600
Qy	178	SerAaPThrIlePheIIeGlnHisAenGlyIleuIleIyValIegIySerValTTPHisArg	197
Ds	601	AGGAGCACCATCTTCATTCAGACCAACGGCTCATCAAGATCGGTGCTGTGGACCGA	660

QY	139	IIepheSeeAaAlaLeuAaRProProthAlaLeuProAaAaAaLeuAaRSeProIle	217
Db	661	ATCTTCTCCAAT-----GACCTTCCAGATGATCTCCGAAGCCCATC	702
QY	218	ATGAAGIaRGUaRGUaLeuAaRProProthAlaLeuProAaAaAaLeuAaRSeProIle	237
Db	703	CGCGTGAAGCAGAGAGAACTTGGAAACCTGCACCTTCTCCCGCAGATATGGAAGAGTG	762
QY	228	AlaAaRGlyThAlaValAaPLePheSeRPhEgIyMeCysAlaLeuGIuMeAlaVal	257
Db	763	GCGCATGGGACCGCTGTGGACATCTTCTCTTGGGATGTGAGCGCTGGAGATGCTGTA	822
QY	258	LeuGIuIIeGIInThAaNGIaVaSPThRaRGVaIThRGUaIuAlaIleAlaAaRGAlaAaRG	277
Db	823	CTGGAAGAAATCCAAACCAATAGGGAGAACCCGGGTGCACAGAGAGAGCCATTCTCGCCGCAAG	882
QY	278	HisSerLeuSeRaSPProAaMeRaRGGuPHeIIeLeuCySaPheuAlaRaGaAPRo	297
Db	883	CACCTGGCTGAAGTACCCCAACATGGGGAGTTCACTTCTTGCTGCTGGCCCGGAGACCT	942
QY	298	AlaAaRGaRGProSeRaIaHisSerLeuAaPheHIAaRGValLeuPheGIuAlaHisSer	317
Db	943	GCCCCCGGCGCTCTGCCACAGCCTCTCTTCCACCGGCTGCTTCCAGAGTGACCTCG	1000
QY	318	LeuLySleuLeuAlaAlaHisCySaPheIIeGIInHIAeGIInTyRLeuMeRProGIuAaRaVal	337
Db	1003	CTGAAGCTCTCGACAGCCCACTGCTTATCCAGACCAAGTACCTCAATGCTTGAGAAATGCG	1060
QY	338	ValGIuGIuLySThTyRAlaMeRaAaPheuHIsAlaValLeuAlaGIuLeuPProAaRGPro	357
Db	1063	GTGGAGAGAAAGACCAAGAGCCATGACCTGCACCGGGTCTTGCGAGGCTTCCCGGCCC	1120
QY	358	ARGaRGProProLeuGIInThPaRGTyRSeRGUaValSeRPhMeGIuAaAaPlySPha	377
Db	1123	CGCAGCGCCCGCTGCTGACGTGGCGTACTGGAAAGTCTCTTCAATGAGCTGGACAAATTC	1180
QY	378	LeuGIuAaPValaARgaNGIyIIeTyRProLeuMeRaAaPheAlaAlaThRaRGProLeu	397
Db	1183	CTGGAGAGATGCAGGAATGAATATCAACCACTGAATGAATCTTGACAGCCACTCGACCCCTG	1240
QY	398	GIyLeuPProAaRGValLeuAlaProProProGIuGIuValGIuLySAlaIyThRProthR	417
Db	1243	GCGCTGCCCCCGTGTCTGACCCCAACCCCGGAGAGGTCCAAAGGCCAAAGACCCCGAGC	1300
QY	418	ProGIuPProPheAaSPSeRGUThRaRGlySValIIeGIInMeGIInCySaAaNGUaRG	437
Db	1303	CCAGAGCCCTTTGACTCTGACACCAAGAAAGTCAATCCAGATCAATGCTGAACCTGGAGAGA	1360
QY	438	SerGIuAaPlySAlaAaRGTrPhIsleuThRLeuLeuValLeuGIuAaPARGLeuHIs	457
Db	1363	AGCGAGGACAAAGCGCGCTGGCATCTGACTGCTTGATGTGCTGGAAAGACCGGGCTGCAC	1420
QY	458	ARGIInLeuThTyRPaPLeuLeuPProThRaSPSeRaIaGIInAaPLeuAlaSeRGUleu	477
Db	1423	CGCGACGTGACCTTACGACTGTCTCCCAACGAGCACCGGCCAGAGACTCGCCTCGAGAGCTC	1480
QY	478	ValHisTyRGlyPheLeuHIsGIuAaSPaARGMeClySleuAlaAlaPheLeuGIuSeR	497
Db	1483	GTGCACATAGCTTCTCTCCACAGAGACGACCGGATGAAGCTGAGCGCCTTCTTGAGAGAC	1540
QY	498	ThRPhLeuLyTyRARGIyThRGInAla507	
Db	1543	ACCTTCTCAAGTACCGTGGAGCCAGGCC1572	


```

275 GCGGCGCAGAGAGAGATCCAGACCGTGTGACAGCTGCTGCTGGAGCACCAG 334
101 AenllleVallyLeuNhiIsyRtYrTTPLeuAerThSerGluAlaCyaaIaaRyValIle 120
335 AACATCGTAAAGTGGACAAGTACCTGGCTGGATACCTCTGAGGCCCTGGAGGCTCATC 394
121 PheIleThGluTyRValSerSerGlySerLeuValGlnPheLeuNhiIsyRtYrTTPLeu 140
395 TTCAATCAGAGATACGTCTCATACAGCAAGCTTCAAGCAATTCCTCAAAAAGACCAAGAG 454
141 AaenHllyValImerAaenHlIaaRgaIaTPrlysaRgTPrCyRThGlnIleLeuSerAla 160
455 AACCAACAGGCGCATGAAACCGCGGCGCTGAAAGCGCTGGAGCAGCAGATTCGTGTGG 514
161 LeuSerPheLeuNhiAlaCySerSerProProlIeIleHiglyAaenLeuThSerAerThr 180
515 CTCACGCTTCTGCACCCCTGCAGCCGCCCAATCATCCACGGGAACTGTACAGCCAGCACCC 574
181 IlePheIleGlnHlIaenGlyLeuIlelyValIleGlySerValTTPHlIaRgIlePheSer 200
575 ATCTTCATTCAGACACAGCGCTCATCAAGATCGCTCCGTGGCAGCAAGATCTTCTCC 634
201 AaenAlaLeuAerProProThraIaLeuProAaPheLeuAerSerProIleAerGlaGlu 220
635 AAT-----GCACTTCCAGATGATCTCCGAAGCCCATCCGCGCTGAG 676
221 ArgGluGluLeuAerAaenLeuNhiIsyRThPheProProGluTyRlyGlyValIaAaerGly 240
677 CAGAGAGAACTTCGGAACCTGCACCTTCTCCCGCCAGAGTATGAGAGGTGGCCGATGG 736
241 ThrAlaValIaerPheSerPheGlyMetCyRAlaLeuGlnMetAlaValLeuGluIle 260
737 ACCGCTGTGGAATCTTCTTCTTGGATGTGTGCGCTGAGATGGCTGTACTGTGAATC 796
261 GlnThraenGlyAerThraRyValThrgluGluAlaIleAlaRgaIaaRghlIserLeu 280
797 CAGACCAATGGGAGACCCGGGCTCAGAGAGGCCATTTGCTCGCCCGAGCACTGCTG 856
281 SerAerProAaenMetAerGluPheIleLeuCyRValLeuAlaRgaAerProAlaAerGly 300
857 AGTGACCCCAACATCGGAGATTCACTCTTCTGCTGGCCGAGACCTCCCGCCGCG 916
301 ProSerAlaHlIserLeuNhiIsyRThraRyValIlePheGluValIleHlIserLeuVal 320
917 CCTCTGTCCACAGCTCTCTTCCACNCGCGTCTTNGAGGTGACACTCGCTGAAGCT 976
320 uLeu 321
977 CCTG 980

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/840,787
FILING DATE: 23-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/518,865
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 2029 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITV03
CLONE: 864683
SEQUENCE DESCRIPTION: SEQ ID NO: 69 :
US-09-840-787-69
US-10-618-941-99 (1-507) x US-09-840-787-69 (1-2029)
Alignment Scores:
Pred. No.: 2,266-175
Score: 1534.50
Percent Similarity: 73.54%
Best Local Similarity: 61.62%
Query Match: 57.47%
DB: 3 Gaps: 6
17 GUAAGGUAAspGluSerGluAspGluSerAerPheGluGlnSerProCyRglYArg 36
240 GAGGAAGAAGAAAGTGAAGATGATGCTGAGATTTGGAAAGTGGCCCTGTGGGCC 299
37 TTPGlnlysaRgaRgGluGlnValaenGlnGlyAaenMetProGlyLeuGlnSerThraPhe 56
300 TGGCAAGAGGCGGAAAGAGTGAATCAACGGAATGTAACAGGTATTCACAGTGCATAC 359
57 leuAlaMetAerThraGluGluGlyValaGluValaITTPAenGluLeuNhiIsyRglYAsp 76
360 CTGGCCATGTATACAGAGAGAGTGTAGAGGTGTGTGGAATGAGAGTACAGTCTCTGA 419
77 ArglysaIaPheAlaHlIseGluGlyValIleGlnThraIlePheGluGlnLeuValLeu 96
420 CGCAAGAACTACAACTCGCAGAGGAAAAAGTTCGTGCTGTGTGTGAATATCTGATTCAA 479
97 ValaerPheProAaenIleVallysaLeuNhiIsyRtYrTTPLeuAerThSerGluAlaCyS 116
480 TTGAGACATCTTAACATTTGTTAAGTTTCAACATATTTGGCGTGAACATTAAGAGAACAG 539
117 AlaRgaValIlePheIleThraGluTyRValSerSerGlySerLeuNhiIsyRglNpheLeu 136
540 GCCAGGTCATTTTATACAGAAATACATGTCATCTGGAGAGTGAAGCAATTTCTGAAG 599
137 lyRThlysaIaenHlIyAlaIaMetAaenAlaRgaIaTPrlysaRgTPrCyRThGln 156
600 AAGACCAAAAAGAACCAAGACGATGAATGAATGAAGGATGAACCTTTGGTGCAACAA 659
157 IleuSerAlaLeuSerPheLeuNhiIsyRtYrTTPLeuAerThSerGluAlaCyS 176
660 ATCTCTGTGCTTAACTGACCTGCACTCTGTGACCCCGCCCATCATCAATGGAACTTG 719

```

177 ThrsrAaPThrIlePheIleGlnHisAnGlyLeuIleValIleGlySerValTrpHis 196
1780 ACCTGTGACATCTTCTCATCCAGACCAACGACATCAATGATGGCTGTG----- 773
197 ArgIlePheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerPro 216
774 -----GCTCCCGACACATCAACATCAT 797
217 IleArgAlaGluArgGluGluLeuArgAsnLeuHisPhePheProProGluTrgIleu 236
798 GTGAAGACTGTGCGAAGACGACGAAGATCTACACTTCTTGGACCCAGATATGAGAA 857
237 ValAlaAsp---GlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMet 255
858 GTCCTAATATGTACACACAGCATGTGACATCTCTTGGCATGTGTGCTCATGTGAGATG 917
256 AlaValLeuGluIleGlnThrAsnGlyAspThrArg---ValThrGluGluAlaIleAla 274
918 GCAGTGTGTGAGATTCAGGGGCAATGAGAGTCTCATATGTGCCACAGAGGCCATCAGC 977
275 ArgAlaArgHisSerLeuSerAspProAsnMetArgGluPheIleLeuCysLeuAla 294
978 AGTGCATCCAGCTTCTAGAAAGACCATTAACAGAGGAGTTCATTCAAAAGTGCCTGAG 1037
295 ArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGlu 314
1038 TCTGAGCTGTGCGACACCAACAGCCAGAGAACTTCTGTTCACCCAGCATTTGTTAA 1097
315 ValHisSerLeuLeuLeuAlaHisCysPheIleGlnIleGlnTrpLeuMetPro 334
1098 GTGCCCTGTGCTCAAACTCTTGGGCCCCCATCTGCATTTGGGACACCAACATGATCCCA 1157
335 GluAsnValValGluGluLeuThrIleValAspLeuHisAlaValLeuAlaGluLeu 354
1158 GAGAACGCTGTAGAGAGATCACCAAAACATGATATAGTACGCGTACGCTGAATC 1217
355 ProArg---ProArgArgProProLeuGlnTrpArgTrpSerGluValSerPheMetGlu 373
1218 CTTGACGAGCACGAGAAAGAACCAAGTTCAGACTTGTACTCAGATTCACAGCTTGAAA 1277
374 LeuAspLeuPheLeuGluAspValArgAsnGlyIleTrpProLeuMetAsnPheAlaAla 393
1278 TTAGTAAATTCCTTGAAGATTCAGGAATGGGATCTATCTCTACAGCCTT----- 1331
394 ThrArgProLeuGlyLeuProArgValLeuAlaProProProGluGluVal----- 410
1332 -----GGGCTGCTCGGCCCCAGCAGCAGCAGAGAGAGAGATCATCATCCT 1379
411 -----GlnValAlaIleThrProThrProGluProPheAspSerGluThrArgIys 427
1380 GTCTGCCCCCTCTGTCAAGCTCCGACCTGAAACAGCTGAGAGTGAAGCTGCGAAG 1439
428 ValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLeuValaArgTrpHisLeuThr 447
1440 GTGGTGTGATGCAATGCACTTGAAGTCGGGAGAGAGAGAGTCAAAACACCATCTGACA 1499
448 LeuLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThrTrpAspLeuLeuProThr 467
1500 CTTCTGCTGAAGTGTGAAGACAACTGAACCGGACCTGAGCTGAGCTGAGCTGAGCAAT 1559
468 AspSerAlaGlnAspLeuAlaSerGluLeuValHisTrpGlyPheLeuHisGluAspAsp 487
1560 GAGAAATATCCCGAGTGTGGCTGAGCTGTGACGCTGGGCTTATTAAGTAGAGGCTGAC 1619
488 ArgMetLeuLeuAlaAlaPheLeuGluSerThrPheLeuIleTrp 502
1620 CAGAGCCGCTTCTCTCTGTAGAAAGACCTTGAACAACTTC 1664

GENERAL INFORMATION:
APPLICANT: KONDA, Goichi
APPLICANT: MATSUDA, Akio
APPLICANT: MURAMATSU, Shuji
APPLICANT: ISHIZAWA, Kenya
TITLE OF INVENTION: STAT6 Activating Gene
FILE REFERENCE: 1254-0207P
CURRENT APPLICATION NUMBER: US/10/153,668
CURRENT FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: US 60/293,172
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/316,031
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: US 60/328,403
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: JP 2001-157043
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: JP 2001-260681
PRIOR FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: JP 2001-313175
PRIOR FILING DATE: 2001-10-10
NUMBER OF SEQ ID NOS: 488
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 321
LENGTH: 2163
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (113)..(1717)
US-10-153-668-321
Alignment Scores:
Pred. No.: 2,5e-175 Length: 2163
Score: 1534.50 Matches: 305
Percent Similarity: 73.54% Conservative: 59
Best Local Similarity: 61.62% Mismatches: 102
Query Match: 57.47% Indels: 29
DB: Gaps: 6
US-10-618-941-99 (1-507) x US-10-153-668-321 (1-2163)
17 GluArgGluAspGluSerGluAspGluSerAspIleLeuGluGluSerProCysGlyArg 36
242 GAGAAAGAAAGAAAGAGATGAGCTGAGATTTTGGAAAGTCCCTGTGGGCGC 301
37 TrpGlnIysArgArgGluGlnValaAsnGlnIleAsnMetProGlyLeuGlnSerThrPhe 56
302 TGGCAGAAAGGCCAGAAAGAGTGAATCAACGGAATGACAGATTAAGACAGTGCATAC 361
57 LeuAlaMetAspThrGluGluGlyValGluValaIleTrpAsnGluLeuHisPheGlyAsp 76
362 CTGGCCATGATATACAGAGAAAGTGTAGAGTGTGTGATGAGTACAGTTCTCTGAA 421
77 ArgIysAlaPheAlaAlaHisGluGluLeuIleGlnTrpValPheGluGlnLeuValLeu 96
422 CGCAAGAACTACAGCTGCAGAGGAAAGGTTGTGTGTGTTGAATATCTGATTCAA 481
97 ValAspHisProAsnIleValIleLeuHisIleTrpTrpLeuAspThrSerGluAlaCys 116
482 TTGAGAGATCTTAACATTTGTAAGTTTCACAAATTTGGGCTGACATTAAGAGAAACAG 541
117 AlaArgValIlePheIleThrGluTrpValSerSerGlySerLeuIleGlnPheLeuIys 136
542 GCCAGGCTCATTTTATTCACAGATATCATGTCTGGAGTCTGAAGCAATTTCTGAAG 601
137 LysThrIysIysAsnHisIleValIleMetAsnAlaArgAlaTrpIysArgTrpCysThrGln 156
602 AAGACCAAAAGAACCAACAGACGATGAATGAAGAGATGAAGCGTTGGTGCACACAA 661
157 IleuSerAlaLeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAsnLeu 176
662 ATCTCTCTGCTTAAGTACTGACTCTCTGTGACCCCCCAATCATCAAGGAACTGT 721

177 ThrSerAspThrIlePheIleGlnHisAsnGlyLeuIleValIleGlySerValTrpHis 196
172 ACCGTGAGACCACTTCTCATCCAGCAACAGCACTATCAAGATTGGCTGTGG----- 775
197 ArgIlePheSerAsnAlaLeuArgProProThrAlaLeuProAspPheLeuArgSerPro 216
776 -----GCTCTGACACTTCAACATCAT 799
217 IleArgAlaGluArgGluGluLeuArgAsnLeuHisPhePheProGluTyrGlyGlu 236
800 GTGAAGACTGTGCGAAGAGCAAGAAATCTACCTCTTGGCAACAGATGTGAGAGA 859
237 ValAlaAsp---GlyThrAlaValAspIlePhePheSerPheGlyMetCysAlaLeuGlnMet 255
860 GTCACTATGTGACAAACAGAGTGCACATCTACTCTTGGCAAGTGTGCATGAGAGATG 919
256 AlaValLeuGluIleGlnThrAsnGlyAspThrArg---ValThrGluGluAlaIleAla 274
920 GCAGTGTGAGATTTCAGGGGCAATGAGAGTCTCTATATGTGCCAAGAGCAAGCATCAGC 979
275 ArgAlaArgHisSerLeuSerAspProAsnMetArgGluPheIleLeuCysCysLeuAla 294
980 AGTGCCTATCCAGCTTCTAAGAGCCCATTAACAGAGGAGTTCATCAAAAGTGCCTGAG 1039
295 ArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGlu 314
1040 TCTGAGCTGTCTGCGACCAACAGCAAGAACTTCTGTCCACCCAGCATTTGTTGAA 1099
315 ValHisSerLeuLeuLeuAlaHisCysPheIleGlnHisGlnTyrLeuMetPro 334
1100 GTGCCCTGCTCAAACTCTTGGCGGCCCTGCTATGTGGACACCAACATGATCCCA 1159
335 GluAsnValValGluGluTyrThrIleValMetAspLeuHisAlaValLeuAlaGluLeu 354
1160 GAGAACGCTCTAGAGAGATCAACAAACATGATACTAGTCCGTATCGGCTGAATTC 1219
355 ProArg---ProArgArgProProLeuGlnTyrPheArgTyrSerGluValSerPheMetGlu 373
1220 CTTGCAGAGCAAGAAAGAACCAAGTTCAGACTTGTACTCTGCACACCACTCTGGAA 1279
374 LeuAspPhePheLeuGluAspValArgAsnGlyIleTyrProLeuMetAsnPheAlaAla 393
1280 TTAGATTAATTCCTTGAATATGTCAAGATGGATTAATCTCTGACACCTTT----- 1333
394 ThrArgProLeuGlyLeuProArgValLeuAlaProProProGluGluVal----- 410
1334 -----GGGCTGCTCGGCCCCCAGACCCAGACCAAGAGAGAGTGCATCAGCT 1381
411 -----GlnIleValAlaIleThrProThrProGluProPheAspSerGluThrArgIle 427
1392 GTCTGCCCCCTCTGTCAAGACTCGACACTCGAACCTGAGCTGAGAGTCCGCAAG 1441
428 ValIleGlnMetGlnCysAsnLeuGluArgSerGluAspIleAlaArgTrpHisLeuThr 447
1442 GTGGTGTGATTCAGTGCACATTAAGTGGTGAAGAGAGAGTCAACACCACTCGAAG 1501
448 LeuLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThrTyrAspLeuLeuProThr 467
1502 CTTTGTGTGAAGTTGAGAGCAAACTGAACCGGCACTGAGCTGTGACTGATGCGCAAT 1561
468 AspSerAlaGlnAspLeuAlaSerGluLeuValHisTyrGlyPheLeuHisGluAspAsp 487
1562 GAGAAATATCCCGAGTGTGGCTGTGCTGTGAGCTGGGCTTCAATTGTAGAGGCTGAG 1621
488 ArgMetCysLeuAlaAlaPheLeuGluSerThrPheLeuIleTyr 502
1622 CAGAACCGGTGACTTCTCTGCTAAGAGACCTTGAACAGATTC 1666

RESULT 13
US-10-887-553A--882
; Sequence 882, Application US/10887553A
; Publication No. US20050085436A1

GENERAL INFORMATION:
APPLICANT: Garza, Dan
TITLE OF INVENTION: Method to treat conditions associated
with insulin signalling dysregulation
FILE REFERENCE: 4-33262
CURRENT APPLICATION NUMBER: US/10/887,553A
PRIOR APPLICATION NUMBER: 60/485,883
PRIOR FILING DATE: 2003-08-07
NUMBER OF SEQ ID NOS: 1208
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 882
LENGTH: 2163
TYPE: DNA
ORGANISM: human
US-10-887-553A-882

Alignment Scores:
Pred. No.: 2,5e-175 Length: 2163
Score: 1534.50 Matches: 305
Percent Similarity: 73.54% Conservative: 59
Best Local Similarity: 61.62% Mismatches: 102
Query Match: 57.47% Indels: 29
DB: Gaps: 6

US-10-618-941-99 (1-507) x US-10-887-553A-882 (1-2163)

17 GluArgGluAspGluSerGluAspGluSerAspIleLeuGluGluSerProCysGlyArg 36
242 GAGGAAGAAAGAAAGATGAATGATGATGATTTGGAAGATCGCCCTGTGGCGCG 301
37 TTPGlnLysArgArgGluGluValAsnGlnLysAsnMetProGlyLeuGlnSerThrPhe 56
302 TGGCAGAGAGCGCGAAGAGGTGAATCAACGGAATGTACACAGATTCAGATGATCAG 361
57 LeuAlaMetAspThrGluGluGluValGluValAlaTrpAsnGluLeuHisPheGlyAsp 76
362 CTGGCCATGATACAGAGAGAGGTGAGGTGTGTGGAATGAGGTACAGTCTCTGAA 421
77 ArgLysAlaPheAlaAlaHisGluGluLysIleGlnThrValPheGluGlnLeuValLeu 96
422 CCGAAGACTACAACTGACAGAGAGAAAGTTCGTCGTGTGATTAATCTGATTCAA 481
97 ValAspHisProAsnIleValLysLeuHisLysTyrTrpLeuAspThrSerGluAlaCys 116
482 TTGAGCATCTTAACATTTGTTAAGTTTCAAAATATTGGGCTGACATTAAGCAAG 541
117 AlaArgValIlePheIleThrGluTyrValSerSerGlySerLeuLysGlnPheLeuLys 136
542 GCCAGGGTCATTTTATACAGAAATACATGATCTGGGAGTCTGAGCAATTTCTGAAG 601
137 LysThrLysValAsnHisLysValAlaMetAsnAlaArgAlaTrpLysAspTrpCysThrGln 156
602 AAGACCAAAAGAACCAACAGACATGAATGAAGAGCATGGAAGGTGGACACACA 661
157 IleLeuSerAlaLeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAsnLeu 176
662 ATCTCTGTGCCCTTAAGTCACTGACATCTGTGACCCCCCATCATCATCGAAGACCG 721
177 ThrSerAspThrIlePheIleGlnHisAsnGlyLeuIleLysIleGlySerValTrpHis 196
722 ACCTGTGACACCATTTCTATCCAGCAACAGACATCAATCAATGAGCTGTG----- 775
197 ArgIlePheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerPro 216
776 -----GCTCTGACACTTCAACATCAT 799
217 IleArgAlaGluArgGluGluLeuArgAsnLeuHisPhePheProGluTyrGlyGlu 236
800 GTGAAGACTGTGCGAAGAGCAAGAAATCTACCTCTTGGCAACAGATGTGAGAGA 859
237 ValAlaAsp---GlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGlnMet 255

Db 860 GTCACTAATGTGCAACAAGCAAGTGGACATCTACTCTTTGGCATGTGTGCACCTGGAGATG 919
Oy 256 AAlaValLeuGluIleGlnThrAsnGlyAspThrArg---ValThrGluGluAlaIleAla 274
Db 920 GCAGTGTCTGGAGATTCACAGGCGCAATGGAGAGTCTCATATGTGTCCACAGAAACCATTCAC 979
Oy 275 ArgAlaArgHisSerLeuSerAspProAsnMetArgGluPheIleLeuCysValLeuAla 294
Db 980 AGTGCATATCCACGCTTCTAGAGACCCATTACAGAGGAGATTCATCAAAAAGTCCCTGCAG 1039
Oy 295 ArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGlu 314
Db 1040 TCTGAGCTGTCTGGACAGCAACAGCCAGAGAACTTCGTCTCCACCCAGCATTTGTTTAA 1099
Oy 315 ValHisSerLeuLeuLeuLeuAlaAlaHisCysPheIleGlnHisGlnIleLeuMetPro 334
Db 1100 GTGGCCCTGGCTCAAACTCTTGCGGCCACCTGATCTGTGGACACCAACATGATCCCA 1159
Oy 335 GluAsnValValGluGluLeuThrIleValMetAspLeuHisValValLeuAlaGluLeu 354
Db 1160 GAGAAACGCTCTAGAGAGATCACCAAAAACATGATACTAGTCCGCTACTGGCTGAATTC 1219
Oy 355 ProArg---ProArgArgProProLeuGlnTrpArgIleSerGluValSerPheMetGlu 373
Db 1220 CCTGCAGAGACCAAGAAAGAACCAAGTCAAGCTTTGTACTCTCAAGTCCACAGCTTGGAA 1279
Oy 374 LeuAspLeuPheLeuGluAspValArgAsnGlyIleIleIleProLeuMetAsnPheAlaAla 393
Db 1280 TTAGATTAATTCCTTGAAGAGATGTACGAATGGATCTATCTCTACAGCCTTT----- 1333
Oy 394 ThrArgProLeuGlyLeuProArgValLeuAlaProProProGluGluVal----- 410
Db 1334 -----GGGCTGCTCTGGCCCCAGCCAGCCAGACAGAGAGAGAGTGCATCACCT 1381
Oy 411 -----GlnValAlaIleThrProThrProGluProPheAspSerGluThrArgGly 427
Db 1382 GTCTGCCCCCTCTGTCAAGCTCCGACACTGAAACCAAGCTGAGAGTGAAGCTCGCAAG 1441
Oy 428 ValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLeuValAlaArgTrpHisLeuThr 447
Db 1442 GTGGTGTGATGCACTGACACATTAGATCGGTGGAGGAGGAGAGTCAAAACACCACTGACA 1501
Oy 448 LeuLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThrIleIleAspLeuLeuProThr 467
Db 1502 CTTCTGTGAAGTGTGAGAGCAAACTGAAACCGGCACCTGAGGTGTGACTGTGCAAT 1561
Oy 468 AspSerAlaGlnAspLeuAlaSerGluLeuValHisIleIleIlePheLeuHisGluAspAsp 487
Db 1562 GAGAAATATCCCGAGTGTGGCTGAGCTGTGGCAGCTGGGCTTCATTATGAGAGGTGAC 1621
Oy 488 ArgMetLeuAlaAlaPheLeuGluSerThrPheLeuIleIleIleIleIleIleIle 502
Db 1622 CAGAGCCGCTGACTTCTCTGTAGAAAGAGACCTTGAAACAAGTTC 1666
RESULT 14
US-10-956-157-1845
Sequence 1845, Application US/10956157
Publication No. US20050118625A1
GENERAL INFORMATION:
APPLICANT: Wyeth
TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
FILE REFERENCE: 031896-043000 (AM 101081)
CURRENT APPLICATION NUMBER: US/10/956,157
NUMBER OF SEQ ID NOS: 2004-10-04
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1845
LENGTH: 2163
TYPE: DNA
ORGANISM: Homo sapiens

US-10-956-157-1845			
Alignment Scores:			
Pred. No.:	2,5e-175	Length:	2163
Score:	1534.50	Conservative:	305
Percent Similarity:	73.54%	Mismatches:	59
Best Local Similarity:	61.62%	Indels:	102
Query Match:	57.47%	Gaps:	6
DB:	9		
US-10-618-941-99 (1-507) x US-10-956-157-1845 (1-2163)			
QY	17	GIUARGIUAERGIUSERGIUAERGIUSERAELLLEUGLGIUSERPROCYGSIYARG	36
DB	242	GAGGAAGAAGAAAGTGAAGATAGTCTGAGATTGTAAGAGAGTCGCCCTGTGGGCGC	3010
QY	37	TRPGILYVATGATGUGLGLVAlaENGInGLYAMMePRoGLYEUENInSERTPhE	56
DB	302	TGGCGAAGAGCGCGAAGAGGTGAATCAACGAATGTACAGAGTGTACACATGCATAC	361
QY	57	LEUAlaMeAERPTGIUGLUGIYAlaIUAIVAlaTPRANGLIUNHISPhEGIYAEPR	76
DB	362	CTGGCGCATGATACAGAGAGAGTGTAGCTGTGTGTGAATGAGGTACAGTTCTGTGA	421
QY	77	ARGIYAlaPheAlaAlaHISGLUGILYUVEGINThValPheGLUGInLeuAlaLeu	96
DB	422	CGCAAGAACTACAACTGCAGAGGAAAGGTTCCGTGCTGTGTGATATCTGATTCAA	481
QY	97	VALAAPHISPRoAENIIEValLYLEUHIISLYETVTRPRLEUAERPTSERGIUAlaCY	116
DB	482	TTGGAGCATCTTAACATGTTGAAGTTTACAAATATTTGGGCTGACATTAAAGAACAG	541
QY	117	AlaARQVAlIIEPHeIIEThRGILUTYValSERSErGLISerLEUySGInPHeLEUyS	136
DB	542	GCGAGGTCATTTTATATACAGAAATACATGTATCTGGAGCTGAAGCAATTTCTGAAG	601
QY	137	LYETHILYLYVAENHIISLYAlaMeAENAlaARGAlaTPRILYAlaTRPCYETHrGIn	156
DB	602	AAGAACAAAAGAACCAAGACATGATGAAGAAAGCATGAGAGGCTTGTCACACAA	661
QY	157	ILELEUSERAlaLEUSERPHeLEUHIISAlaCYSErPRoPTIIEIENISGLYANLEU	176
DB	662	ATCTCTTGCCCTTAAGCTACCTGACCTTCGTGACCCCCCATCATCTCATGGGAACCTG	721
QY	177	ThISEAERPTThIIEPHeIIEGLNIISANGLIYUAlaIETVIEIIEGLYSErValTRPHIS	196
DB	722	ACCTGTGACACCATTTATCTTATCCAGACAAAGGACATCAAGATTGGCTCTGTG----	775
QY	197	ArgIIEPHeSERANAlaLEUAETRPRoThrAlaLEUPRoAEPVAlaEUAERSErPRo	216
DB	776	-----GCTCTGCACCTATCAACATCAT	799
QY	217	ILEARGAlaGLIUAERGILUGILEUAARGANLEUHIISERPHePRoPROGLIUTYrGILU	236
DB	800	GCGAAGACTTGTGCGAAGAGCGAAGATCTTAACTTTTGGACACCAAGATGAGAA	859
QY	237	VALAlaAER---GLYThAlaValAERPIIEPHeSERPHeGLIYMeCYAlaLEUGIUMET	255
DB	860	GTCACATAATGTGACAAACAGCACTGACATCTTACTCTTGGACATGTGTCTACTGGAGATG	919
QY	256	AlaValLEUGInIIEGINThrASNGIYAErThrARG---ValThrGLIUGIAlaIeAla	274
DB	920	GCAAGGTCTGAGATTCAGAGGCAATGGAAGTCTCATATGTGTCCACAGAAAGGCATCAG	979
QY	275	ARGAlaARGHIISERLEUSERAErPRoAENMeCAlaUGILUPHeIIELEUCYCYVLEUAla	294
DB	980	AGTGCATACCACTTCTGAAGACCCATTACAGAGGAGTTCATTCMAAGTCCCTGCGAG	103
QY	295	ARGAERPRoAlaARGARGPRoSErThIISERLEUyUPHeNIEAlaGLAUErPHeGLI	314
DB	1040	TCTGAGCTGTCTCGAGAACCAACAGCCCAAGAACTTCTGTGTCCACCCSACCATTTGTTGA	109
QY	315	ValHISERLEUyVLEUAlaAlaHISCYVrPHeIIEGLNIISGLIUTYrLEUeMePRo	334

Db 1100 GTGCCCCCTGCTCAAACTCTTGGCGCCACTGATTTGGGACACCAACCATGATGCCA 1159
Qy 335 GUAaerValValGluGluGluSerThrValaMetAspLeuHisAlaValLeuAlaGluLeu 354
Db 1160 GAGAAAGCTCTAGAGAGATGACCAAAAACATGAGATATAGTCCCGATGCTGGTGAATC 1219
Qy 355 ProArg---ProArgArgProProLeuGlnTrpArgTyrSerGluValSerPheMetGlu 373
Db 1220 CTTGACGAGACCAAGAAAGAACCAAGTTCAGACTTTGTACTCTCAGTCACCAAGCTTGGAA 1279
Qy 374 LeuAspLysPheLeuGluAspValArgAsnGlyIleTyrProLeuMetAsnPheAlaAla 393
Db 1280 TTAGATTAATTCCTTGAAGATGTCAGAAATGGGATGATATCTCTGACAGCTTT----- 1333
Qy 394 ThrArgProLeuGluLysProArgValLeuAlaProProProGluGluVal----- 410
Db 1334 -----GGGCTGCTCTGGCCCGCCAGCAAGCCACAGCAGAGAGAGGATCATCAGCT 1381
Qy 411 -----GlnLysAlaLysThrProThrProGluProPheAspSerGluThrArgLys 427
Db 1382 GTTCGCCCCCTCTCTCAAGACTCCGACACTGAACCAAGCTGAGTGAAGCTGCGCAAG 1441
Qy 428 ValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLysAlaArgTrpHisLeuThr 447
Db 1442 GTGGTGCTGATGAGTGAACCAATGATGCTGGTGGAGAGAGAGATCAACACCACTGACA 1501
Qy 448 LeuLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThrTyrAspLeuLeuProThr 467
Db 1502 CTTCGCTCAAGTTGGAGAGCAAACTGAACCGGCACCTGAGCTGATGATGCCAAT 1561
Qy 468 AspSerAlaGlnAspLeuLysSerGluLeuValHisTyrGlyPheLeuHisGluAspAsp 487
Db 1562 GAGAAATATCCCGAAGTTGCGGCTGAGCTGGTGGACCTGGGCTTCAATTAGTGAAGCTGAC 1621
Qy 488 ArgMetLysLeuAlaAlaPheLeuGluSerThrPheLeuLysTyr 502
Db 1622 CAGAGCCGGTGTACTTCTGCTAGAAAGACCTTGAACAAGTTC 1666
RESULT 15
US-09-764-875-151
; Sequence 151, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PAMM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 151
; LENGTH: 2225
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-875-151
Alignment Scores:
Pred. No.: 2,61e-175 Length: 2225
Score: 1534.50 Matches: 305
Percent Similarity: 73.54% Conservative: 59
Best Local Similarity: 61.62% Mismatches: 102
Query Match: 57.47% Indels: 29
DB: 3 Gaps: 6
US-10-618-941-99 (1-507) x US-09-764-875-151 (1-2225)
Qy 17 GUAaerGluAspGluSerGluAspGluSerAspIleLeuGlnGluSerProCysGlyArg 36
Db 270 GAGGAAGAAAGAAAGTGAAGATGATGATGATTTGGAAAGATGCGCCCTGTGGGCGC 329
Qy 37 TrpGlnLysArgArgGluGlnValaHisGlnGlnLysMetProGlyLeuGlnSerThrPhe 56

Db 330 TGGCAAAAGAGCGAAGAAAGGATGATCAACGGAATGTACAGATTTTACAGTGCATAC 389
Qy 57 LeuAlaMetAspThrGluGluGlyValGluValaTrpAsnGluLeuHisPheGlyAsp 76
Db 390 CTGGCCATGATGATCAAGAGAAAGGTGTAGAGTTGTGTGGAAATGAGATGATCTCTGAA 449
Qy 77 ArgLysAlaPheAlaAlaHisGluGluLysIleGlnThrValaPheGluGlnLeuValLeu 96
Db 450 CGGAAGAACTACAAAGCTGACAGAGAAAGTTGCTGTGTGTATATATCATGATTCAA 509
Qy 97 ValAspHisProAsnIleValLysLeuHisLysTyrTrpLeuAspHisSerGluAlaCys 116
Db 510 TTGAGACATCTTAACATTTGTAATTTTCAAAATATTGGCTGACATTAAAGAAACAG 569
Qy 117 AlaArgValaIlePheIleThrGluTyrValaSerSerGlySerLeuLysGlnPheLeuLys 136
Db 570 GCGAGGTCATTTTATATCAGAAATATCATGTATCTGGGAGTGTGAAGCAATTTCTGAAG 629
Qy 137 LysThrLysLysAsnHisLysValaMetAsnAlaArgAlaTrpLysArgTrpCysThrGln 156
Db 630 AAGACCAAAAGAAACCAAGAAAGATGAAGCAATGAAGCATGGAACGTTGTGCACAA 689
Qy 157 IleLeuSerAlaLeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAsnLeu 176
Db 690 ATCTCTTCTGCTTAAGCTTACCTGACCTCTGTGACCCCTCATATCCATGGAACCTG 749
Qy 177 ThrSerAspThrIlePheIleGlnHisAsnGlyLeuIleLysIleGlySerValTrpHis 196
Db 750 ACCTGAGACACCATTTTCATCCAGCAACAGGATCATGAATTTGGCTGTGG----- 803
Qy 197 ArgIlePheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerPro 216
Db 804 -----GCTCTGACACTTCAACAAATCAT 827
Qy 217 IleArgAlaGluArgGluGluLeuArgAsnLeuHisPhePheProProGluTyrGlyGlu 236
Db 828 GTGAAGACTTGTGAGAAAGAGAGAAAGATCTACACTTTTGGACACCAAGATGTGGAGAA 887
Qy 237 ValAlaAsp---GlyThrAlaValaAspIlePheSerPheGlyMetCysAlaLeuGlnMet 255
Db 888 GTCACTAATGTGACAACAGCACTGACATCTACTCTTTGGATGATGTGTCACTGGAGATG 947
Qy 256 AlaValLeuGluIleGlnThrAsnGlyAspThrArg---ValaThrGluGluAlaIleAla 274
Db 948 GCAGTGTGAGATTCAGGGCAATGAGAGTCTCATATGTCCACAGAGAACCATACAGC 1007
Qy 275 ArgAlaArgHisSerLeuSerAspProAsnMetArgGlnPheIleLeuCysCysLeuAla 294
Db 1008 AGTGCATTCACAGCTTCTAGAAAGACCATTTACAGAGGAGATTCATTCAAAAGTCCGAG 1067
Qy 295 ArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValaLeuPheGlu 314
Db 1068 TCTGAGCTGTGCTCGAGAACCAACAGCAGAGAACTTGTGTCCACCCAGCATTTTGA 1127
Qy 315 ValHisSerLeuLysLeuLeuAlaAlaHisCysPheIleGlnHisGlnTyrLeuMetPro 334
Db 1128 GTGCCCTGCTCAAACTCTTGGCGCCCACTGATTTGGGACACCAACACATGATCCCA 1187
Qy 335 GUAaerValValGluGluLysThrValaMetAspLeuHisAlaValLeuAlaGluLeu 354
Db 1188 GAGAAAGCTCTAGAGAGATGCCAAAACATGATATAGTGGCCGTGCTGGTGAATC 1247
Qy 355 ProArg---ProArgArgProProLeuGlnTrpArgTyrSerGluValSerPheMetGlu 373
Db 1248 CTTGACGAGACCAAGAAAGAACCAAGTTCAGACTTTGTACTCTCAGTCACCAAGCTTGGAA 1307
Qy 374 LeuAspLysPheLeuGluAspValaArgAsnGlyIleTyrProLeuMetAsnPheAlaAla 393
Db 1308 TTAGATTAATTCCTTGAAGATGTCAGAAATGGGATGATATCTCTGACAGCTTT----- 1361
Qy 394 ThrArgProLeuGluLysProArgValLeuAlaProProProGluGluVala----- 410

```
Db      1362 -----GGGCTGCTCGGCCAGCCAGCCAGAGAGAGTGCATCACCCT 1409
Qy      411 -----GlnlysaAlaIythrProthrProGluProPheaspSergIuThrArgIys 427
Db      1410 GTCGTGCCCTCTGTCAAGACTCGACACCTGAACCACTGAGGTGAGACTCGCAAG 1469
Qy      428 ValIleGlnMetGlnCysAsnLeuGluAArgSergIuAAspIysAlaArgTrpHisLeuThr 447
Db      1470 GTGGTGTGATGCAGTGCACATTGATCGGTGGAGGAGGAGTCAAAACACCCTGACA 1529
Qy      448 LeuLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThrTyrAspLeuLeuProThr 467
Db      1530 CTTCTGTGAAGTTGGAGGACAACTGAACCGGCACCTGAGCTGTGATGCCAAT 1589
Qy      468 AspSerAlaGlnAspLeuAlaSerGluLeuValHisTyrGlyPheLeuHisGluAspAsp 487
Db      1590 GAGATATATCCCGAGTTGGCGGCTGAGCTGGGCTTCATTATGAGGCTGAC 1649
Qy      488 ArgMetLysLeuAlaAlaPheLeuGluSerThrPheLeuLysTyr 502
Db      1650 CAGAGCCGGTTGACTTCTGTCTAGAGAGACCTTGAAACAAGTTC 1694
```

Search completed: January 14, 2006, 03:03:27
Job time : 996 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 13, 2006, 20:06:05 ; Search time 742 Seconds
(without alignments)
4553.907 Million cell updates/sec

Title: US-10-618-941-99
Perfect score: 2670
Sequence: 1 MAAPBPAPRARERERERED.....RMKLAFLPSTFLKXGTQA 507

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model -DEV=xlp
-O=/cgn2.1/USPTO_epool_p/US10619941/runat_12012006_070203_14258/app_query.fasta_1.647
-DB=N Geneseq -QFMT=fastap -SUFFIX=ing -MIMMARCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human0.cdi -LIST=45
-DOCALLIN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10618941@cgn_1_1.1096@runat_12012006_070203_14258 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2670	100.0	3765	12	ADJ96576 Human N1m
2	2612	97.8	4048	12	ADL13324 Human N1e
3	2584	96.8	3837	13	ADK52130 Plant ful
4	2584	96.8	3838	13	ADK52108 Plant ful

5	2584	96.8	3846	13	ADK52133	Adk52133 Plant ful
6	2531.5	94.8	3128	13	ADK54234	Adk54234 Plant ful
7	2499	93.6	3147	14	ADV97751	Adv97751 CDNA sequ
8	2409	90.2	3304	4	AAK4681	Aak4681 Novel pro
9	2409	90.2	3304	12	ADI29379	Adi29379 Human MAR
10	2368	88.7	3538	13	ADV15041	Adv15041 CDNA diff
11	2368	88.7	3538	14	AEA61757	Aea61757 Human nuc
12	2201	82.4	1501	4	AAK51824	Aak51824 Human pol
13	1617	60.6	981	5	AAK29899	Aak29899 Human h1s
14	1534.5	57.5	2141	5	AAH14504	Aah14504 Human CDN
15	1534.5	57.5	2158	6	ABA93734	Ab93734 Human bra
16	1534.5	57.5	2163	10	AAK4684	Aak4684 Novel pro
17	1534.5	57.5	2163	10	ADG10731	Adg10731 Human STA
18	1534.5	57.5	2163	12	ADI29382	Adi29382 Human MAR
19	1534.5	57.5	2163	14	ADZ49553	Adz49553 Human MAR
20	1534.5	57.5	2163	14	AEA61756	Aea61756 Human nuc
21	1534.5	57.5	2225	12	ABK43561	Abk43561 DNA encod
22	1534.5	57.5	2225	12	AD153948	Ad153948 CDNA encod
23	1529.5	57.3	2159	5	AAH14231	Aah14231 Human CDN
24	1529.5	57.3	2159	5	AAH78063	Aah78063 Nucleotid
25	1522.5	57.0	2178	10	ADG10843	Adg10843 Human STA
26	1508.5	56.5	2241	2	AAV59121	Aav59121 Nucleotid
27	1502.5	56.3	1931	3	AAK98922	Aac98922 Human pan
28	1360.5	51.0	2328	4	AAK4682	Aak4682 Novel pro
29	1360.5	51.0	2328	12	ADI29380	Adi29380 Mouse MAR
30	1174.5	44.0	2803	4	AB102019	Ab102019 Drosophil
31	1170	43.8	2923	8	ABE80409	Ab280409 Human pro
32	1106.5	41.4	1568	4	AAK4685	Aak4685 Novel pro
33	1106.5	41.4	1568	12	ADI29383	Adi29383 Mouse MAR
34	1087.5	40.7	1606	5	AAK64747	Aac64747 Secreted
35	1022.5	38.3	5655	4	AB102018	Ab102018 Drosophil
36	965.5	36.2	1980	12	ADK58748	Adk58748 DNA encod
37	955	35.8	1065	4	AAK27150	Aak27150 CDNA encod
38	955	35.8	1065	4	ABK4365	Abk4365 DNA encod
39	955	35.8	1065	10	ADB93328	Adb93328 Human CDN
40	955	35.8	1065	12	ADI54252	Adi54252 CDNA encod
41	810.5	30.4	2002	12	ADK15884	Adk15884 4 synthe
42	756	28.3	538	10	ADI22646	Adi22646 Human liv
43	756	28.3	910	4	AAK27174	Aak27174 CDNA encod
44	756	28.3	910	4	ABK43562	Abk43562 DNA encod
45	756	28.3	910	10	ADB93352	Adb93352 Human CDN

ALIGNMENTS

RESULT 1	ADJ96576	standard; DNA; 3765 BP.
ID	ADJ96576	
XX	ADJ96576;	
AC	ADJ96576;	
DT	06-MAY-2004 (first entry)	
XX		
DE	Human N1m-A related protein kinase NRPB2 DNA SeqID 33.	
XX		
KM	gene; ds; kinase; human; tyrosine protein kinase;	
KM	serine/threonine protein kinase; PTK; STK; gene therapy; cancer;	
KM	immune-related disease; cardiovascular disease; brain;	
KM	neural associated disease; metabolic; inflammatory disorder;	
KM	cytosolic; neuroprotective; immunomodulator; antiinflammatory;	
KM	N1m-A related protein kinase; NRPB2.	
XX		
OS	Homo sapiens.	
OS	53.	
PN	W0200406838-A2.	
XX		
PD	22-JAN-2004.	
XX		
PF	15-JUL-2003; 2003WO-US021730.	
XX		
PR	15-JUL-2002; 2002US-039633P.	
XX		

PA (SUGC-) SUGEN INC.

XX Whyte D, Manning G, Caenepeel S;

XX WPI: 2004-122753/12.

DR P-PSDB; ADJ96642.

XX New nucleic acid molecule encoding a kinase polypeptide, useful for
PT preparing a composition for treating diseases or disorders, e.g., cancer,
PT or neurological, immunological or inflammatory disorders.

XX Example 1; SEQ ID NO 33; 366pp; English.

CC This invention relates to a novel isolated, enriched or purified nucleic
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC to human tyrosine and serine/threonine protein kinases (PK's and STK's),
CC as well as protein kinase-like enzymes. The present invention describes
CC screening methods to identify agonists, antagonists and antibodies that
CC can be used to modulate the activity or function of the mammalian kinase
CC enzymes. As such, these compositions can be used for gene therapy
CC purposes to treat diseases or disorders including cancer, immune-related
CC diseases, cardiovascular diseases, brain or neuronal associated disease,
CC metabolic and inflammatory diseases. Accordingly, they exhibit
CC cytoprotective, neuroprotective, immunomodulator and anti-inflammatory
CC activities. This polynucleotide sequence is a human kinase DNA sequence
CC of the invention.

SQ Sequence 3765 BP; 801 A; 1168 C; 1040 G; 756 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4,04e-259	Length:	3765
Score:	2670.00	Matches:	507
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12	Indels:	0
		Gaps:	0

US-10-618-941-99 (1-507) x ADJ96576 (1-3765)

QY 1 MetAlaAlaProGluProAlaProAlaProAlaArgAlaArgGluArgGluArgGluArg 20
Db 282 ATGGCGGCGCCCGGAGCCCGCGCGCGAGCGCGCGCGCGGAAACGGGAGCGGAGGAGC 341
QY 21 GluSerGluAapGluSerAspIleuGluGluSerProGlyArgTrpGlnArg 40
Db 342 GAGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 401
QY 41 ArgGluGlnValAsnGlnGlyAsnMetProGlyLeuGlnSerThrPheLeuAlaMetAsp 60
Db 402 CGGAGCGAGGTAACCAAGGAAATGCCAGGAGCTTCAGAGCACTTCCTAGCCATGAGC 461
QY 61 ThrGluGluGlyValGluValAlaTrpAsnGluLeuHisPheGlyAspArgIleAspHe 80
Db 462 ACGGAGGAGGCGGTGAGAGGTGTGAGAACGAGCTCCACTTCGAGAGACGAGAGGCTTC 521
QY 81 AlaAlaHisGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 100
Db 522 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 581
QY 101 AsnIleValIleLeuHisIleuSerGlySerLeuGluGlnPheLeuIleuValIle 120
Db 582 AACATCGGAGATTGCAAGATCTGGCTGAGTACTCTGAGGCTTCGCGAGGAGTATC 641
QY 121 PheIleThrGluTrpValSerSerGlySerLeuGluGlnPheLeuIleuValIle 140
Db 642 TTCATCAGAGATGATGTCATCAGGAGCTCAAGCAATTCCTCAAAAAGCAAGAG 701
QY 141 AsnHisIleValMetAsnAlaArgAlaTrpIleuArgTrpGlyThrGlnIleuSerAla 160
Db 702 AACCAAGGCGCATGAAAGCCCGGCGCTGGAAGCCCTGTGTGACCAAGATCTCTGCTCG 761
QY 161 LeuSerPheLeuHisAlaCysSerProPheIleHisGlyAsnLeuThrSerAspThr 180

Db 762 CTCAGCTTCCTGACAGCGCTGAGCCCCCAATCATCCAGGGAACCTGACCAGGACACC 821
QY 181 IlePheIleGlnHisAsnGluLeuIleuSerIleGlySerValIleTrpHisArgIlePheSer 200
Db 822 ATCTTCATTCAGCAACAAGGCTTCATCAAGATCGGCTCCGCTGTCAGCCGATCTTCCTCC 881
QY 201 AsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIleArgAlaGlu 220
Db 882 AATGCTCTGGCCCTTCACAGCACTTCAGATATCTCCGAAGCCCATCCGCGCTGAG 941
QY 221 ArgGluGluLeuArgAsnLeuHisPhePheProProGluTrpGlyGluValAlaAspGly 240
Db 942 CGAGAGGAATTCGGAACCTGCACTTCCTCCCAAGATGTAGAGAGTGGCGATGG 1001
QY 241 ThrAlaValAlaPhePheSerPheGlyMetCysAlaLeuGluMetAlaValLeuGluIle 260
Db 1002 ACCGCTGTGACATCTTCCTTGGAGATGTGTCGCTGAGATGCGCTGATCTGGAATC 1061
QY 261 GlnThrAsnGlyAapTrpArgValThrGluGluAlaIleAlaArgAlaArgHisSerLeu 280
Db 1062 CAGACCAATGGAGGACACCGGAGTCAAGAGAGGCCATTGCTCCGCGAGGACATCGCTG 1121
QY 281 SerAspProAsnMetArgGluPheIleuGlyCysIleuAlaArgAspProAlaArgArg 300
Db 1122 AGTACCCCAACATGCGGAGATTCATCTTGTGCTGCTGCGCCGAGACCTGCGCGCGG 1181
QY 301 ProSerAlaHisSerLeuLeuPheHisArgValIleuPheGluValHisSerLeuIleu 320
Db 1182 CCGTCTGCCACACGCTCTTCACCGCGCTGCTTCGAGGTGCATCTGCTCAAGGTC 1241
QY 321 LeuAlaAlaHisCysPheIleGlnHisGlnTrpLeuMetProGluAsnValAlaGluGlu 340
Db 1242 CTGGAGCCCACTCTCTTCATCAGACACAGTACCTCATGCTTGAGATGTGTGAGAGAG 1301
QY 341 LysThrIleValMetAspLeuHisAlaValIleuAlaGluLeuProArgProArgArgPro 360
Db 1302 AAGACCAAGGCGCATGAGCTGCAAGCGGTCTTGCGAGGCTTCCCGGCGCCGAGGCGCC 1361
QY 361 ProLeuGlnTrpArgIleTrpSerGluValIleSerPheMetGluLeuAspArgIlePheGlu 380
Db 1362 CCGTGTGAGTGGCGGTCTTCGGAAGTCTCTTCATGAGCTGAGCAAAATTCCTGGAGGAT 1421
QY 381 ValArgAsnGlyIleTrpProLeuMetAsnPheAlaIleThrAspProLeuGlyLeuPro 400
Db 1422 GTCAAGAAATGAACTTACCACTATGAACTTTCAGGCACTTCGACCTCTGGGCTGCC 1481
QY 401 ArgValIleuAlaProProProGluGluValGluIleuValIleuValIleuValIleu 420
Db 1482 CGTGTGCTGGCCCAACCCCGGAGAGGTCCAAAAGGCCAAGACCCGAGCCAGAGGCC 1541
QY 421 PheAspSerGluTrpArgIleValIleGlnMetGlnCysAsnLeuGluArgSerGluAsp 440
Db 1542 TTTGACTCTGAGACAGAAAGTCAATCAGATGCAAGTGCACCTGAGAGAAAGGAGAGAC 1601
QY 441 LysAlaIleArgTrpHisIleuThrLeuLeuValIleuGluAspArgLeuHisArgGlnLeu 460
Db 1602 AAGGCGGCTGGCAATCTACTCTCTGTGCTGAGAACCCGCTGACCGGAGCTG 1661
QY 461 ThrTrpAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuValHisIleTrp 480
Db 1662 ACCTACAGACCTGCTCCCAAGGAGCAGGCCCAAGAGACCTCGCTGGAGCTGTGACATAT 1721
QY 481 GlyPheLeuHisGluAspAspArgMetCysIleuAlaIlePheLeuGluSerThrPheLeu 500
Db 1722 GGTCTCTCCCAAGAGGAGCAGCCGATGAAGCTGCGCTTCCTGAGAGACACTTCTCTC 1781
QY 501 LysTrpArgGlyThrGlnAla 507
Db 1782 AAGTACCGTGGAGCCAGGCC 1802

RESULT 2
ADL13324
ID ADL13324 standard; cDNA; 4048 BP.

XX ADL13324;
 AC 06-MAY-2004 (first entry)
 XX
 DT Human steroid-induced C3A liver cell cDNA #1053.
 XX
 DE ss; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;
 XX steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
 XX
 OS Homo sapiens.
 XX
 PN US673549-B1.
 XX
 PD 06-JAN-2004.
 XX
 PF 12-OCT-2001; 2001US-00976594.
 XX
 PR 12-OCT-2000; 2000US-0240409P.
 XX
 PA (INCYTE) INCYTE CORP.
 XX
 PI Furness LM, Buchbinder JL,
 XX
 DR WPI; 2004-068610/07.
 XX
 PT Combination useful for preparing a composition for treating liver
 PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
 PT comprises cDNAs that are differentially expressed in response to steroid
 PT treatment.
 XX
 PS Claim 1; SEQ ID NO 1053; 141bp; English.
 XX
 CC The invention relates to a combination comprising cDNAs that are
 CC differentially expressed in response to steroid treatment. Also included
 CC are the following: a high throughput method for using a cDNA to detect
 CC differential expression of nucleic acids in a sample; and a high
 CC throughput method of screening molecules or compounds to identify a
 CC ligand that specifically binds a cDNA. The sample is from a subject with
 CC Wilson disease and comprises of a standard defines a stage of that
 CC disease. The high throughput method of screening molecules or compounds
 CC to identify a ligand that specifically binds a cDNA comprises: combining
 CC the combination with molecules or compounds under conditions to allow
 CC specific binding; and detecting specific binding between each cDNA and at
 CC least one molecule or compound. The molecules or compounds are regulatory
 CC proteins. The combination is useful for preparing a composition for
 CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
 CC or hepatitis. The present sequence represents a human cDNA which is
 CC differentially expressed in steroid-induced C3A liver cells. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html.
 CC
 XX
 SQ Sequence 4048 BP; 862 A; 1229 C; 1080 G; 877 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 3,32e-253 Length: 4048
 Score: 2612.00 Matches: 501
 Percent Similarity: 98.62% Conservative: 0
 Best Local Similarity: 98.62% Mismatches: 0
 Query Match: 97.83% Indels: 7
 DB: 12 Gaps: 1
 US-10-618-941-99 (1-507) x ADL13324 (1-4048)

QY 41 ArgGluGlnValAsnGlnGlyAsnMetProGlyLeuGlnSerThrPheLeuAlaMetAsp 60
 DB 402 CCGAGCAGGAGTAAACCAAGGAAACATGCGAGGCTTCAAGACACTTCTAGCCATGAC 461
 QY 61 ThrGluGlnGlyValGluValValTrpAsnGluLeuHisPheGlyAspArgValAlaPhe 80
 DB 462 ACGAGAGAGGGGGTGAAGGTGGTGTGGAAAGAGCTTCACTTGGAGACAGAAAGGCTTC 521
 QY 81 AlaAlaHisGluGlnGlyValLeuGlnThrValPheGluGlnLeuValLeuValAspHisPro 100
 DB 522 GCGGGGCAACGAGAAAGAAATCCAGACCGGTGTCCAGACCTGATGCTGTGAGCAACCCG 581
 QY 101 AsnIleValIleLeuHisValSerTyrTrpLeuAspThrSerGluAlaCysAlaArgValIle 120
 DB 582 AACATCGTAAGTTCACAACTGAGTCTGGATACCTCTGAGGCTTGGCCGAGGATCATC 641
 QY 121 PheIleThrGluTyrValSerSerGlySerLeuGlyGlnPheLeuValLeuValThrIleVal 140
 DB 642 TTCATCAGAGATGATGTCATCAGGACGCTCAAGCAATTCCTCAAAAAGCAAGAG 701
 QY 141 AsnHisValValMetAsnAlaArgAlaTrpIleValGlyTrpCysThrGlnIleLeuSerAla 160
 DB 702 AACCAAGGCGCATGAACCGCCGGGCTGGAAGCGCTGTGACGACGATCCTGTCTGG 761
 QY 161 LeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAsnLeuThrSerAspThr 180
 DB 762 CTCAGCTTCTCAGCGCTGACGCCCCCAATCATCCAGGAACTTGACGACGACAC 821
 QY 181 IlePheIleGlnHisAsnGlyLeuValLeuValIleGlySerValTrpHisAspGlyPheSer 200
 DB 822 ATCTTCATTCAGACACACGCGCTCATCAAGATCGGCTCGGTGGACCGAATCTTCTTC 881
 QY 201 AsnAlaLeuValGlyProProThrAlaLeuProAspAspLeuValSerProIleArgAlaGlu 220
 DB 882 AAT-----GCACTTCAGATGATCTCCGAAGCCCATCCGCGCTGAG 923
 QY 221 ArgGluGlnLeuValArgAsnLeuHisPhePheProProGlyTyrGlyGlnValAlaAspGly 240
 DB 924 CAGAGGAACTTCGAAACCTGCACTTCTTCCCCCAAGTAAGAGAGTGGCCATGTGG 983
 QY 241 ThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaValLeuGluIle 260
 DB 984 ACCGTGTGACATCTTCTCTTGGAGTGTGGCTGGAGATGGCTGTATCTGGAATTC 1043
 QY 261 GlnThrAsnGlyAspThrArgValThrGluGlnAlaIleAlaArgAlaArgHisSerLeu 280
 DB 1044 CAGACCAATGGGGAACCCGGGTGCACAGAGGACCATTTGCTGGCCAGGACATCGCTG 1103
 QY 281 SerAspProAsnMetArgGluPheIleLeuCysCysLeuAlaArgAspProAlaArgArg 300
 DB 1104 AGTGACCCCAACATGCGGAGTTCACTTGTCTCTGCGCCCGGAGCCCTGCGCGG 1163
 QY 301 ProSerAlaHisSerLeuLeuPheHisAspArgValLeuPheGluValHisSerLeuValLeu 320
 DB 1164 CCTCTGCCCAACGCTCTCTTCCACCGCGTCTCTTCAAGTGTGACATCTCGTGAAGCTC 1223
 QY 321 LeuAlaAlaHisCysPheIleGlnHisGlnTyrLeuMetProGluAsnValValGluGlu 340
 DB 1224 CTGGAGGCCCACTGTTTCATCAGACCAAGTACTTCATGCTGAGAAATGTGGTAGAGAG 1283
 QY 341 LysThrIleValMetAspLeuHisAlaValAlaLeuAlaGluLeuProArgProArgArgPro 360
 DB 1284 AAGACCAAGGCGATGACCTGACCGGCTTGGGAGGCTTCCCGGCGCCGACGCGCC 1343
 QY 361 ProLeuGlnTrpArgTyrSerGluValSerPheMetGluLeuAspIlePheLeuGluAsp 380
 DB 1344 CCGCTGCAAGTGGCTTCTCGGAAGTCTCTTCAAGAGCTGACAAATCTCTGAAGAGAT 1403
 QY 381 ValArgAsnGlyIleTyrProLeuMetAsnPheAlaAlaThrArgProLeuGluIleuPro 400
 DB 1404 GTCAAGAAATGGAATCTACCACTGATGAACCTTGGACGACATCGACCCCTGGGGCTGCC 1463
 QY 401 ArgValLeuAlaProProProGluGlnValGlnValAlaLysThrProThrProGluPro 420

DB 1464 CGTGTGCTGGCCCAACCCCGAGAGAGGTCCAAAAGGCCAAGACCCCGAGCCGAGACC 1523
QY 420 OPheAspSerGIuThrAglYValIIGImeGInCYsAnLeuGIuAgtSerGIuAs 440
DB 1524 CTTGACCTCTGAGACAGAAAGGTATCCAGATGACGTGCAACTGGAGAGAGCGAAGA 1583
QY 440 pLyAlaRgtTrpHisLeuThrLeuLeuValLeuGIuAspAglYLeuHisArgGInLe 460
DB 1584 CAAGCGCGCTGGCACTCCTCTCTCTGCTGGAAGACCGGCTGCAACCGGCACT 1643
QY 460 UThrYrAspLeuLeuProThrAspSerAglInAspLeuAAserGIuLeuValHisTy 480
DB 1644 GACCTACGACCTGCTCCCAACGAGACAGCGCCCAAGACCTCCGCTCGAGCTCTGCTACTA 1703
QY 480 rGIYpHeLeuHisGluAspAspArgMetIYsLeuValAlaAphLeuGIuSerThrPheLe 500
DB 1704 TGGCTTCTCTCCACGAGACGACCGGATGAGCTGCGCTTCTCTGAGAGACACTTCTCT 1763
QY 500 ULyEThrArgGlyThrGlnAla 507
DB 1764 CAAGTACCGTGGAGACCGAGCC 1785
RESULT 3
ADXS2130
ID ADXS2130 standard; cDNA; 3837 BP.
AC ADXS2130;
XX 21-APR-2005 (first entry)
DE Plant full length insert polynucleotide seqid 26870.
XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX unidentified.
OS US2004034888-A1.
PN 19-FEB-2004.
PD 28-APR-2003; 2003US-00425114.
PF 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TAB/) TABASKA J E.
PA (CAO/) CAO Y.
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JB, Cao Y;
XX WPI; 2004-180133/17.
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX Claim 1; SEQ ID NO 26870; 1583; English.
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence

CC available in electronic form from the US patent office at
CC ftp.segdata.uspo.gov/sequence.html?docid:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
SQ Sequence 3837 BP; 799 A; 1195 C; 1084 G; 759 T; 0 U; 0 Other;
XX
XX
Alignment Scores:
Pred. No.: 2,096-250 Length: 3837
Score: 2584.00 Matches: 501
Percent Similarity: 88.05% Conservative: 0
Best Local Similarity: 88.05% Mismatches: 0
Query Match: 96.78% Indels: 68
DB: 13 Gaps: 2
US-10-618-941-99 (1-507) x ADXS2130 (1-3837)
QY 1 MetAlaAspProGIuProAlaProArgArgAlaArgGIuArgGIuArgGIuAsp 20
DB 231 ATGGCGGCCCCGAGACCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 290
QY 21 GluSerGIuAspGIuSerAspIleLeuGIuGIuSerProCYsGIYArgTrpGInLYArg 40
DB 291 GAGAGCAGAGACGAGACGACATCTCTGAGAAAGCCGCTGTGCTCGCTGCAAAAGCGA 350
QY 41 ArgGIuGIuValAengInGIYAsnMerProGIYLeuGIuSerThrPheLeuAlaMetAsp 60
DB 351 CGGAGAGGTAAACCAAGGAAACATGCCAGGCTTGAAGACACTTCTAGCCATGGAC 410
QY 61 ThrGIuGIuGIYValGIuValValITrPAsnGIuLeuHisPheGIYAspArgLYAlaPhe 80
DB 411 ACGAGAGAGGGGAGGTAGGTGTGTGGAACGAGCTCCACTTCGAGACAGGAAGGCTTC 470
QY 81 AlaAlaHisGIuGIuValIleGIuThrValPheGIuGIuLeuValIleAspIlePro 100
DB 471 GCGGCGCAGAGAGAGATCCAGACCGTGTTCGAGCAGCTGGGTGTGTGACACCCG 530
QY 101 AsnIleValIYsLeuHisLYrITrPLeuAspThreGIuAlaCYsAlaArgValIle 120
DB 531 AACATCGTGAAGTTCACCAAGTACTGGCTGTGATACCTCTAGAGGCTTCGCGAGGTATC 590
QY 121 PheIleThrGIuTrpValIleSerSerGIYSerIleuYsGInPheLeuLYsTrpThrLYsLYs 140
DB 591 TTCATTCAGAGTACGTGTCTCTACGSCAGCTCTCAAGCAATTCCTCAAAAAGACCAAGAG 650
QY 141 AsnHisLYsAlaMetAlaAlaArg----- 148
DB 651 AACCAAGAGGCATGAAAGCCCGGATGGGAGCGGCGTGGGAGCAGCAGCGGAGACAG 710
QY 148 ----- 148
DB 711 ACGGAGTTGGGAGACCTCGGAGACTGGAGATGGAGGGGTGCGCGGCGGCTCGAGCA 770
QY 148 ----- 148
DB 771 GGGGCTGGGAGAGATCGGGGCGGAGCTCCGAGAGGCCAGCGCTCTCTGCGCCAC 830
QY 149 -----AlaTrpLYsArgTrpCYsTrpGInIleu 158
DB 831 GACCGAGAGATGTGTGCTCGCGCGGAGGCTGGAAGCGCTGTGTCAAGCATATCTGT 890

QY 159 SerAlaLeuSerPheLeuH1sAlaCySerProProl1e11eH1sGlyLeuLeuThrSer 178
 DB 891 TCTGGGCTCAGCTTCTTGACCGCTGACGGCCCAATCATCCAGGAACTGACACAC 950
 QY 179 AapThrIlePheH1eGlnH1sAaNGlyLeuIleyleIleGlySerValItrpH1sArgIle 198
 DB 951 GACACCATCTTCATCAGACAAACGGCTCATCAAGATGGCTCCGCTGGCCACCAATC 1010
 QY 199 PheSerAspAlaLeuArgProProthraIleuProAspAspLeuArgSerProl1eArg 218
 DB 1011 TTCTCCAAAT-----GCACTTCAGAAATGATCTCGAAGCCCAATCCCGC 1052
 QY 219 AlAGluArgGluGluLeuArgLeuH1sPhePheProProGluItrGlyGluValAla 238
 DB 1053 GCTGAGCGAGAGAACTTGGAACTGCACTTCTCCCAAGATGAGAGAGTGGCC 1112
 QY 239 AapGlyThrAlaValAapIlePheSerPheGlyMetCysAlaLeuGluMetAlaValLeu 258
 DB 1113 GATGGGACCGCTGTGGACATCTTCTCTTGGGATGTGTGGCTGGAGATGGCTGTACTG 1172
 QY 259 GluIleGlnThrAsnGlyAspThrArgValIthrGluGluAlaIleAlaArgAlaArgHis 278
 DB 1173 GAAATCCAGACCAATGGAGGACACCCGGGTCAAGAGAGGCCATGCTCGCCAGGAC 1232
 QY 279 SerLeuSerAspProAspMetArgGluPheIleLeuCysCysLeuAlaArgAspProlAla 298
 DB 1233 TGGCTGAGTGAACCCCAACATGGGGAGTTCTTCTTGGCTGCTGGCCGGACCTGTGC 1292
 QY 299 ArgArgProSerAlaH1sSerLeuLeuPheH1sArgValIleuPheGluValH1sSerLeu 318
 DB 1293 CCGCGGCCCTGTGCCACAGCCCTCTTCCACCGGTGCTGTGGAGGTGACCTGTGCTG 1352
 QY 319 LysLeuLeuAlaAlaH1sCysPheH1eGlnH1sGlnItrLeuMetProGluLeuValVal 338
 DB 1353 AAGCTCTGTGGACGCCCACTGCTTCATCCAGACCACTATCTGATGATGTGTGTG 1412
 QY 339 GluGluLeuThrIleAlaMetAspLeuH1sAlaValIleuAlaGluLeuProArgProArg 358
 DB 1413 GAGGAGAGAACCAAGGCCATGACCTGACCGCGGTCTTGGCGAGCTTCCCGGCCCC 1472
 QY 359 ArgProProLeuGlnItrArgItrSerGluValIserPheMetGluLeuAspLysPheLeu 378
 DB 1473 AGGCCCGCGCTGACAGGCGGTACTCGGAAGTCTCTTATGAGAGCTGACAAATTCTCG 1532
 QY 379 GluAapValArgAsnGlyIleItrProLeuMetAspPheAlaAlaThrArgProLeuGly 398
 DB 1533 GAGGATGTCAAGAAATGAAATCTCAACCACTGATGAATCTTGACGCCACTGACCCCTGGGG 1592
 QY 399 LeuProArgValIleuAlaProProProGluGluValGlnIleAlaGlyThrProThrPro 418
 DB 1593 CTGCCCGGTGTCTGGCCCCCGGAGGAGGTCCAAAGGCCCAAGCCCGACGCCA 1652
 QY 419 GluProPheAspSerGluThrArgIleValIleGlnMetGlnCysAsnLeuGluArgSer 438
 DB 1653 GAGCCCTTGAATCTGAGACCAAGAAAGTCAATCCAGATCAAGTGCACCTGAGAGAGAAC 1712
 QY 439 GluAapValArgItrpH1sLeuThrIleLeuLeuValIleGluGluAspArgIleH1sArg 458
 DB 1713 GAGGACAAGCGCGCTGGCATCTCACTGTCTGTGTCTGAGAGACCGGCTGACCGG 1772
 QY 459 GlnLeuThrItrAspLeuLeuProThrArgSerAlaGlnAspLeuAlaSerGluLeuVal 478
 DB 1773 CAGCTGACCTACGACCTGCTCCCAACGAGACACCGCCAGAGACTGCGCTCGAGCTCGG 1832
 QY 479 HisItrGlyPheLeuH1eGlnAspAspArgMetLysLeuAlaAlaPheLeuGluSerThr 498
 DB 1893 CACTATGGCTTCTCTCAAGAGACCAACCGGATGAAGCTGCGGCTTCTGAGAGACACC 1892
 QY 499 PheLeuLysItrArgGlyIthrGlnAla 507
 DB 1893 TTCTCTAAGTACCTGTGGACCAAGGCC 1919
 RESULT 4

ADX52108
 ID ADX52108 standard; cDNA; 3838 BP.
 AC ADX52108;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Plant full length insert polynucleotide seqid 26848.
 XX
 KW plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content; gene; ss.
 XX
 OS Unidentified.
 XX
 EN US2004034888-A1.
 XX
 PD 19-FEB-2004.
 XX
 PF 28-APR-2003; 2003US-00425114.
 XX
 PR 06-MAY-1999; 99US-00304517.
 PR 05-NOV-2001; 2001US-00985678.
 XX
 PA (LITU/) LITU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABAS/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 XX
 DR WPI; 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 26848; 15pp; English.
 XX
 CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.nepco.gov/sequence.html?docID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.
 XX
 SQ Sequence 3838 BP; 803 A; 1194 C; 1077 G; 764 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,096-250 Length: 3838
 Score: 2584.00 Matches: 501
 Percent Similarity: 88.05% Conservative: 0
 Best Local Similarity: 88.05% Mismatches: 0

Query Match:	96.78%	Indels:	68
DB:	13	Gaps:	2
US-10-618-941-99 (1-507) x ADX52108 (1-3838)			

OY	1	MetAlaAlaProGluProAlaProArgArgAlaArgGluArgGluArgGluArgGluArg	20
Db	208	ATGGGGCCCCCGAACCCGGCCGGAGCGGGCCGGGAACGGGAACGGAGCGGGAGAC	267
OY	21	GluSerGluAspGluSerAspIleLeuGluGluSerProCysGlyValTropGlnIlyAspArg	40
Db	268	GAGACCGAGAGACGAGAGACCACTCTCGAGAGAAACCCGTGTGTCTGCTGGCAAAAGCA	327
OY	41	ArgGluGlnValAsnGlnGlyAsnMetProGlyLeuGlnSerThrPheIleuAlaMetAsp	60
Db	328	CGGGAGCAGGTAAACCAAGGAGAACTGCAAGGGCTTCAGAGCACTTCCTAGCCATGAGC	387
OY	61	ThrGluGluGluValGluValValTyrPheGlnIleuIlePheGlyAspAspArgIlePhe	80
Db	388	ACGGAGAGAGGGGTAGAGGTGTGTGTGAACGAGCTCCACTTCGAGAGACGAAAGGCTTTC	447
OY	81	AlaAlaHISGluGlnIlyIleGlnThrValPheGluGlnLeuValLeuValAspHisPro	100
Db	448	CGGGCGCAGAGAGAGAGATCCAGAACCGTGTTCGAGAGACTGTGTGTGTGACCAACCG	507
OY	101	AsnIleValIlyLeuHISlyIlyTyrTyrLeuAspThrSerGluAlaCysAlaArgValIle	120
Db	508	AACATCTGTGAAGTTGCACAACTACTCGCTGGATACCTCTGAGAGCCCTGGCGCGAGGTCTATC	567
OY	121	PheIleThrGlnIlyValSerSerGlySerIleuIyGlnPheLeuIlyIlyThrIlyIys	140
Db	568	TTCAATCAGAGATAGCTGTATCATACAGACAGCCCTCAAGCAATCTCTCAAAAGACCAAGAG	627
OY	141	AsnHISlyAlaMetAsnAlaArg-----	148
Db	628	AACCAAGGCGCATGAGCGCCGGGATAGGGAGCGGGCTGGGGAGCAAGGGGACAG	687
OY	148	-----	148
Db	688	ACGGGGTTGGGGCAGCCTCGGGAGCTGGGATGTGAGGGGGTGCACGCGGCTTCGACA	747
OY	148	-----	148
Db	748	GGGGCTGGCGAGAGATCGGGGGCGGGCTCCGAGGCGCAGCCGCTTCCTGCGCCAC	807
OY	149	-----AlaTyrIlyAspTyrCysThrGlnIleu	158
Db	808	CGACCGACGAGTCTGTCGCTCCGCCGCCAGCGCTGTGAAGCGCTGTGACAGCAGATCTCTG	867
OY	159	SerAlaIleuSerPheLeuHISAlaCysSerProProlIleHISGlyAsnIleuThrSer	178
Db	868	TTTGGGCTAGCTCTTCGACGCGCTGACAGCCCCCAATCATACAGGAACTGACAGC	927
OY	179	AspThrIlePheIleGlnHISAsnGlyLeuIleIlyIleGlySerValTyrHisArgIle	198
Db	928	GACACCATCTTCATTACAGACAAAGCGCTCATCAAGATCGCTCCGTGTGTGAGACGAATC	987
OY	199	PheSerAsnAlaLeuAspProThrAlaLeuProAspAspLeuAspSerProIleArg	218
Db	988	TTCTCTCAAT-----GCACCTTCAGATGATCTCCGAAGCCCTCATTCGC	1028
OY	219	AlaGluArgGluGlnLeuArgAsnIleuHISphePheProProGluTyrGlyValAla	238
Db	1030	GCTGAGCGAGAGAACTTCGAAACTGCACTTCTCCCGCCAGAGATGAGAGAGGTGGC	1088
OY	239	AspGlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaValIleu	258
Db	1090	GATGGGACCGCTGTGGACATCTTCTCTTGGAGATGTGTGGCTGTGAGATGCTGTACTG	1144
OY	259	GluIleGlnThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArgHIS	278
Db	1150	GAATATCCAGACCAATGGGGACACCGGGGTGCACAGAGAGGGCCATGTCTGCGCCACGAGAC	1200

QY	279	SeriesSerAaPProAmeMetArgGluPheIleuCySvSvLeuAaIArgAaPProAa	298
Db	1210	TCGCTGAAGTACCCCAACATCGGGAGAGTCACTCTTGCTGCTGGCCGGGACCTTGC	1263
QY	229	ArgAaPProAaIaAaIaSerLeuLeuPheIaArgVaIleuPheGluVaIHiSerLeu	318
Db	1270	CGCGGCGCTTGCCACAGCTCTCTTCCACGCGGCTCTTGAGAGTCACTGCTG	1322
QY	319	LySLeuLeuAaIaAaIaHiSvPheIleGlnHiSvGlnTyLeuMetProGluAaVaIaI	338
Db	1330	AAAGCTCCGGAGGCCACTGCTTATCCAGACCAAGTACCTCAAGCTCGAAGATGTGTG	1389
QY	339	GluGluTyThrIlyVaIaMetAaPLeuHiSvAaVaIleuAaGluLeuProArgProArg	358
Db	1390	GAGGAGAAAGAACAGCCATCGACTGACCGCGCTTGCGGAGAGCTTCCCGGCGCGCG	1449
QY	359	ArgProPProAaGlnTyPArgTySerGluVaIaSerPheMetGluAaPLeuPheLeu	378
Db	1450	AGGCCCCCGCTGACGTGGCGGTACTCGAATCTCTTCAAGAGCTGGAAGAAATTCCTG	1509
QY	379	GluAaPLeuAaIArgAaGlnIlyIleTyProLeuMetAaPheAaIaIaThrArgProLeuGly	398
Db	1510	GAGAGTGCAGGAATGTGAATTAACCACTAAGAACTTTGAGACGACCTCGACCCCTGGGG	1565
QY	399	LeuPProArgVaIleuAaIaProProProGluGluVaIaGlnIlyAaIyThrProThrPro	418
Db	1570	CTGCCCCGCTGTGCTGGCCCCACCCCGGAGAGGTCCAAAGGAGCCAGACCCCGACGCCA	1629
QY	419	GluPProPheAaPserGluThrArgIlyVaIleGlnMetGlnCyAaPLeuGluAaSer	438
Db	1630	GAGCCCTTTGACTCTGAGAACCAAGGATCTCAAGATGCACTGACACTTGGAAGAGAGC	1689
QY	439	GluAaPLeuAaIArgTyPHeIleThrLeuLeuLeuVaIleuGluAaPArgLeuHiAArg	458
Db	1690	GAGAGCAAGGCGCGCTGCATCTCACTCTGTTGCTGCTGGAAGACCGCTGCACCGGG	1749
QY	459	GlnLeuThrTyPLeuAaPLeuProThrAaPserAaIaGlnAaPLeuAaIaSerGluLeuVaI	478
Db	1750	CAGCTGACCTTAAGACCTGCTCCCAACGAGACGCGCCAGGACCTCGCTCGAGCTCGTG	1809
QY	479	HisTyThrIlyPheLeuHiSvGluAaPArgMetIlySvLeuAaIaIaPheLeuGluSerThr	498
Db	1810	CACATAGCTTCTCTCCACAGAGAGACACCGGATGAAGCTGCTGCTCTTGAGAGACACC	1869
QY	499	PheLeuIlyTyPArgGlyThGlnAaI	507
Db	1870	TTCTCTCAAGTACCTGTGGAGCCAGGCC	1896
RESULT 5			
AD52133			
XX	ID	AD52133	standard; cDNA; 3846 BP.
XX	XX	AD52133;	
XX	XX	21-APR-2005	(first entry)
XX	DE	Plant full length insert polynucleotide seqid 26873.	
XX	XX	plant protectant; plant growth regulant; gene therapy; plant;	
XX	XX	recombinant DNA construct; physical array; plant breeding marker;	
XX	XX	cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;	
XX	XX	extreme osmotic condition; pathogen tolerance; pest tolerance;	
XX	XX	growth rate; cell cycle pathway; disease resistance;	
XX	XX	galactomannan production; lignin production; plant growth regulator;	
XX	XX	yield; plant growth; plant development; seed oil; protein yield;	
XX	XX	protein content; gene; ss.	
XX	OS	unidentified.	
XX	PN	US2004034888-A1.	
XX	PD	19-FEB-2004.	

XX		28-APR-2003; 2003US-00425114.
XX		06-MAY-1999; 99US-00304517.
BR		05-NOV-2001; 2001US-00985678.
XX		(LIU/J.) LIU J.
PA	(ZHOU/) ZHOU Y.	
PA	(KOVA/) KOVALIC D K.	
PA	(SCRE/) SCREEN S E.	
PA	(TABA/) TABASKA J E.	
PA	(CAOV/) CAO Y.	
PI	Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;	
XX	WPI; 2004-180133/17.	
DR		
XX		
PT	New recombinant DNA construct, useful for improving plant tolerance to	
PT	cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or	
PT	pests, for conferring increased resistance to plant disease, or for	
PT	improving yield.	
PS	Claim 1, SEQ ID NO 26873; 15bp; English.	
XX		
CC	The invention describes a recombinant DNA construct comprising a	
CC	polynucleotide consisting of a sequence encoding an amino acid sequence	
CC	available in electronic form from the US patent office at	
CC	ftp.segdate.uspo.gov/sequence.html?DocID:2004034888. The polynucleotide	
CC	of the invention are also useful in physical arrays of molecules and as	
CC	plant breeding markers. The recombinant DNA construct is useful for	
CC	improving plant tolerance to cold, heat, drought, herbicides, extreme	
CC	osmotic conditions, pathogens or pests, for manipulating growth rate in	
CC	plant cells by modification of the cell cycle pathway, for conferring	
CC	increased resistance to plant disease, for producing galactomannan,	
CC	lignin or plant growth regulators, for increasing the rate of homologous	
CC	recombination in plants, for improving yield by modification of	
CC	photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake	
CC	or by providing improved plant growth and development under at least one	
CC	stress condition or for modifying seed oil or protein yield and/or	
CC	content. This sequence represents a plant full length insert	
CC	polynucleotide that can be used in the recombinant DNA construct of the	
CC	invention.	
XX		
SQ	Sequence 3846 BP, 805 A, 1197 C, 1081 G, 763 T, 0 U, 0 Other;	
	Alignment Scores:	
	Pred. No.:	2,1e-250 Length: 3846
	Score:	2584.00 Matches: 501
	Percent Similarity:	88.05% Conservative: 0
	Best Local Similarity:	88.05% Mismatches: 0
	Query Match:	96.78% Indels: 68
	DB:	Gaps: 2
US-10-618-941-99 (1-507) x ADX52133 (1-3846)		
OY	1 MetAlaAlaProGluIbuProAlaProArgArGaLaArgGuaArgGuaArgGuaArgGuaAsp	20
Db	216 ATGGCGGCCCGCAGACCGGCAGCGCGCGCGCGCGCGGAACGGGAGCGGAGCGGAGCGAC	275
OY	21 GluSerGluAbnGluSerAsp11LeuGluGluGluSerProCyEGLYAGTTPGlnLysArg	40
Db	276 GAAGAACAAGGACGAAGCAACATCTCGAGGAAACCCGTGTGTGCTGTGGCAAAACCGA	335
OY	41 ArgGluGlnValAbnGlnGlyAbnMetProGlyLeuGlnSerThrPheLeuAlaMetAsp	60
Db	336 CGGAGAGAGTTAACCAAGGAACATGCCAGSGCTTCAGACACCTTCTTAGCATGTGAC	395
OY	61 ThrGluGluGluValGluValValTrpAbnGluLeuHisPheGlyAspArgLYeAlaPhe	80
Db	396 ACCGAGAGAGGGGTAGAGGTGTGTGGAACGAGCTCCACTTCGAGACACGGAAGGCTTC	455
OY	81 AlaAlaHisGluGluLysValIleGlnThrValPheGluGlnLysValIleValAlaPheHisPro	100
Db	456 GCCGCGCACGAAGGAAGATCCAAGACCGTGTTCCAGACAGCTGTGTGTGTGTGACCAACCG	515

QY	101	Asn1Val1LysLeuH1slystyrtlrPheAspThrSerGluAlaCysAlaArgVal11e	120
Db	516	AACATCGTAAGTGTGACAAAGTACTGAGTACTCTGAGGCTTGCCGAGGGTCATC	575
QY	121	Phe1IethrgIuTyrValSerSerGlySerLeuLysGlnPheLeuLysblyThrLys	140
Db	576	TTCAATCAGAGTACTGTGCATCAGGCAAGCTTCAGCAATTCCTCAAAAAGCAAGAG	635
QY	141	AsnH1slyValAmcAsnAlaArg-----	148
Db	636	AAACCAAGGCGCATGAAGCCCGGGTATAGGGAGCGGGCTGGGGCAGCCAGGGACAGG	695
QY	148	-----	148
Db	696	ACGGGGTTGGGGCAGCCTCGGGGAATGGATGTGAAGGGGTGTCCCGAGGCTCGAGACA	755
QY	148	-----	148
Db	756	GGGGCTGGGGCGAAGATCGGGGGGGGCTCCGAGGCCAGCGGCTCTCTCGGGCCACC	815
QY	149	-----AlaTrpLysArgTrpCysThrGlnIleLeu	158
Db	816	CGACCGAGCGAGTCTGGGTCCGCGCGCAGGCGCTGGAAGCGTGGTGACAGCAATCTCG	875
QY	159	SerAlaLeuSerPheLeuH1sAlaCysSerProProlle1H1sGlybAsnLeuThrSer	178
Db	876	TCTGGCGCTCAGGTTCTCTCAGCAGCCCTGCAGCCCCCAATCATCAGGGGAACCTGCAGCAG	935
QY	179	AspThr1IlePhe1IeGlnH1sAsnGlyLeuIleLys1IeGlySerValTrpH1sArg1Ie	198
Db	936	GACACCATCTTCATTACAGACAAAGCGCTCATCAAGATCGGCTCGTGGACCCGAATC	995
QY	199	PheSerAsnAlaLeuArgProThrAlaLeuProAspAspLeuArgSerPro1IeaArg	218
Db	996	TTCTCCAT-----GCACTTCAAGATGATCTCGAAGCCCATCCGC	103
QY	219	AlaGluArgGluGluLeuArgAsnLeuH1sPhePheProProGluTyrglyValAla	238
Db	1038	GCTGAGCAGAGAACTTCGGAACTTCGACATTCCTCCCGCCAGATATGAGAAGGTGGCC	109
QY	239	AspGlyThrAlaValAsp1IlePheSerPheGlyMetCysAlaLeuGluMetAlaValLeu	258
Db	1098	GATGGAGCCGCTGTGACATCTCTCCCTTGGAGTGTGGCCCTGAGATGGCTGTACG	115
QY	259	Glu1IeGlnTrpAsnGlyAspThrArgValThrGluGluAla1IeAlaArgAlaArgH1s	278
Db	1158	GAATTCACAGCAATGGGAGACCCGGGTACAGAGAGGCCATGTCTGCGCCAGGGCAC	121
QY	279	SerLeuSerAspProAsnMetArgGluPhe1IeLeuCysCysLeuAlaArgAspPro1a	298
Db	1218	TGCGTAGTAGCCCAACATGCGGAGTTCATCTTGTGCTGCTGGCCCGGAGCCTGCG	127
QY	299	ArgArgProSerAlaH1sSerLeuLeuPheH1sArgValIleuPheGluValH1sSerLeu	318
Db	1278	CGCCGGCCCTCTGGCCACAGCCTCTCTTCCACCGCGTGTCTTGAGGGTCATCGCTG	133
QY	319	Lys1IeLeuAlaAlaH1sCysPhe1IeGlnH1sGlnTyrlleuMetProGluAsnVal1a	338
Db	1338	AAGCTCCCGAGGCCACTGTCTTATCCAGACCCAGTACCTCATGCTCGAAGATGGTG	139
QY	339	GluGluArgThrLysAlaMetAspLeuH1sAlaValIleuAlaGluLeuProArgProArg	358
Db	1398	GAGAGGAAGACCAAGCGCATGGAAGCTGACCGGTCTTGGCGAGCTTCCCGGCCGCC	145
QY	359	ArgProProLeuGlnTrpArgTyrsGluValSerPheMetGluLeuAspLysPheLeu	378
Db	1458	AGGCCCCCGCTGACGTGGCGGTACTCGAAGATCTCTTCATGAGACTGGACAAATTCTG	151
QY	379	GluAspValaArgAsnGly1IeTyrrProleuMetAsnPheAlaAlaThrArgProLeuGly	398
Db	1518	GAGATGTCAAGAAATGAAATCTACCACTGTAGAACTTTGACGACCACTGACCCCTGGGG	157

OY		339	LeuPfoAglvAlleuAlaProForProGluGluValGlnLysValLysThrProThrPro	418
Db		1578	CTGCCCGGTGCCTGGCCCCCAACCCCGAGGATGCCAAAGCCAAAGCCCGCAGCCCA	1637
OY		419	GluPcPhaArgSerGluThrArgLysValIleGlnMetGlnCyAsnLeuGlnArgSer	438
Db		1658	GAGCCCTTTGAATCTCTGAGACCAGAAAGGTCAITCCAGATGCACTGGAGAGAAC	1697
OY		439	GluARLyValIAArgTPrhIAleuthrLysLeuLeuValLeuGluARgLyLeuNI	458
Db		1698	GAGGACAAGGCGCGCTGGCATCTCACTGCTTCGTGTGTGGAGAACGGGCTGACCGG	1757
OY		459	GlnLeuThrTyArgLeuLeuProThrArgSerIaGlnARLyValIAserGluLeuVal	478
Db		1758	CAGCTGACCTTAGACACTGCTGCCAACGGCACACCGCCAGACCTGGACGTGCTGTG	1817
OY		479	HISrTGlyRPhLeuNIeGluARARArGMetLysLeuAlaIaPhLeuGluSerThr	498
Db		1818	CAGTAATGGCTTCTCCACAGAGACGACCGGATGAAGCTGGCCGCTTCGTGAGAGAC	1877
RESULT 6				
ADX54234				
ID	ADX54234	standard; cDNA; 3128 BP.		
XX	ADX54234;			
AC				
XX	21-APR-2005	(first entry)		
DT				
DE		Plant full length insert polynucleotide seqid 28974.		
XX				
KW		plant protectant; plant growth regulant; gene therapy; plant;		
KM		recombinant DNA construct; physical array; plant breeding marker;		
KW		cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;		
KM		extreme osmotic condition; pathogen tolerance; pest tolerance;		
KW		growth rate; cell cycle pathway; disease resistance;		
KM		galactomannan production; lignin production; plant growth regulator;		
KW		yield; plant growth; plant development; seed oil; protein yield;		
KM		protein content; gene; ss.		
XX				
OS		Unidentified.		
PN	US2004034888-A1.			
XX				
PD	19-FEB-2004.			
XX				
PF	28-APR-2003; 2003US-00425114.			
XX				
PR	06-MAY-1999; 99US-00304517.			
XX	05-NOV-2001; 2001US-00985678.			
PA	(LIUJ/) LIU J.			
PA	(ZHOU/) ZHOU Y.			
PA	(KOVA/) KOVALIC D K.			
PA	(SCRE/) SCREEN S E.			
PA	(TABAS/) TABASKA J E.			
PA	(CAOY/) CAO Y.			
PI	Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y,			
DR	MP1; 2004-180133/17.			
XX				
PT	New recombinant DNA construct, useful for improving plant tolerance to			
PT	cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or			
PT	pests, for conferring increased resistance to plant disease, or for			
XX	improving yield.			
ES	Claim 1; SEQ ID NO 28974; 15pp; English.			
XX				
CC	The invention describes a recombinant DNA construct comprising a			

	polynucleotide consisting of a sequence encoding an amino acid sequence available in electronic form from the US patent office at ftp.segdata.uspto.gov/sequence.html?docID=2004034888. The polynucleotide of the invention are also useful in physical arrays of molecules and as plant breeding markers. The recombinant DNA construct is useful for improving plant tolerance to cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or pests, for manipulating growth rate in plant cells by modification of the cell cycle pathway, for conferring increased resistance to plant disease, for producing galactomannan, lignin or plant growth regulators, for increasing the rate of homologous recombination in plants, for improving yield by modification of photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake or by providing improved plant growth and development under at least one stress condition or for modifying seed oil or protein yield and/or CC content. This sequence represents a plant full length insert
CC	polynucleotide that can be used in the recombinant DNA construct of the invention.
XX	
SQ	Sequence 3128 BP, 617 A; 973 C; 966 G; 572 T; 0 U; 0 Other;
	Alignment Scores:
Pred. No.:	3,17e-245 Length: 3128
Score:	2531.50 Matches: 499
Percent Similarity:	83.47% Conservative: 1
Best Local Similarity:	83.31% Mismatches: 1
Query Match:	94.81% Indels: 99
DB:	Gaps: 3
US-10-618-941-99 (1-507) x ADX54234 (1-3128)	
QY	1 MetAlaAlaProGluPrCoAlaPProArgrAlaArgLuarGluarGluarGluarP 20
Db	120 ATGGCGGCCCGAGACCGGCAGCGCGCGGCCGGAAACGGAGCGGAGCGGAGCAC 179
QY	21 GluSerGluarPgluSerAspIleLeuGlusUserProcGlyAArgTrpGlnyBAry 40
Db	180 GAGAACCGAGAGAGAGAGCATTCTTGAGGAAGACCCTGTGTGTGTGGCAAAAGCCA 239
QY	41 ArgGluGlnValaAnngInglyASandetProGlyLeuGlnSertPhelenuAmAcAp 60
Db	240 CGGAGCAGCTTAAACAAGGAAATTCGACGGCTTCAGAGCATCTTCTTAGCCATGAC 239
QY	61 ThrGluGluGlyValGluValValITrpAnngInleuNHISphEGLyAAsrArglyAlaPhe 80
Db	300 ACGGAGGAGGGGTGAGAGTGGTGTGTGAACGAGCTCACTTGGAGACAGAAAGGCGCTTC 359
QY	81 AlaAlaHISgluGlnUlyLeIGlnThrValAPheGluGlnInleuValleuValasPIsP 100
Db	360 GCGGCGCAGAGAAAMAATCCAGACCGGTGTTCAGACAGCTGTGTGTGACCAACCCG 419
QY	101 AsnIleValLyBleuHISlyETyrTrpLeuAspTrsSerGluAlaCYsaAlarVallie 120
Db	420 AACATCGTAAGACTTCCACAAAGTACTGGCTGATACCTCTGAGAGCCTGCGCGAGGTCAATC 479
QY	121 PheIleThrgIuTyrValSerSerGIyserLeuLygInPheLeuLyAlsyThrlxyBlys 140
Db	480 TTCATCAcAGAgtATcGTGTCAATCAAGCAGCGCTCAAGCAATTCCTCAAHAAGACCAAGAAG 539
QY	141 AsnHISlyValaMetAsnAlaArg----- 148
Db	600 ACGGSGTTGGGCGACGCTTCGGGGACTGGGATGTGAAGGGGGTGCCTCGGCGCTCGGACA 659
QY	148 ----- 148
Db	660 GAGGCTGGGCGAGATGGCGGGCGGGGCTCCGAGGCGCAGCGCGCTCTCTGCGCCAC 719
QY	149 -----AlatTrpLYAsrTrpCySerThrgInlleu 158
Db	720 CGACCGCAGAGATCTGTGCTCGCGCCGACAGCGCTGTGAAGCGCTGTGTGCAcCGAGATCTGTG 779

QY 159 SerAlaLeuSerPheLeuH1sAlaCYsSerProIle1leH1sG1yAseLeuThrSer 178
 Db 780 TCTGCCTCAGCTTCTTGCAGCCTTGCAGCCCTCCCAATCATCCAGGAACTGACACAG 839
 QY 179 AapThrIlePheH1sG1nH1sAeng1yLeu1leY1sG1ySerValTrrH1sArg1le 198
 Db 840 GACACCATCTTCAATCAGACCAACGGCTCATCAATGAGGCTCGGTGGACCGAATC 899
 QY 199 PheSerAsnAlaLeuArgProPrrH1sAlaLeuProAspAspLeuArgSerProIleArg 218
 Db 900 TTCTCCAAAT-----GCACCTTCAGATGATCTCCGAAAGCCCATCCCC 941
 QY 219 AlAG1yArgG1yG1yLeuLeuArgAseLeuH1sPhePheProG1yTrrG1yG1yAla 238
 Db 942 GCTGACGAGAGAACTTGGAACTTGCACCTTTCCTCCCAAGATAGAGAGGTGGCC 1001
 QY 239 AapG1yThrAlaValAsp1lePheSerPheG1yMetCYaAlaLeuG1yMetAlaValLeu 258
 Db 1002 GATGGGACCGCTGTGGACATCTTCTCTTGGGATGTGTGGCTGGAGATGGCTGTACTG 1061
 QY 259 Glu1leG1nThrAeng1yAspThrArgValThrGluGluAla1leAlaArgAlaArgH1s 278
 Db 1062 GAAATTCAGACCAATGGGACACCCGGGTACAGAGAGGCGCATTCCTCGCCAGGCGAC 1121
 QY 279 SerLeuSerAspProAsnMeCArgG1yPheH1sLeuCYeCYeLeuAlaArgAspProAla 298
 Db 1122 TGCCTGAGTGAACCCCAACATGGCGAGGTTCATCTTGTGCTGCTGGCCCGGACCTTGCC 1181
 QY 299 ArgArgProSerAlaH1sSerLeuLeuPheH1sArgValLeuPheG1yValH1sSerLeu 318
 Db 1182 CCGCGGCCCTGTGCCACAGCCTCCTCTTCCACCGGTCTCTTCCAGGTGACCTGCTG 1241
 QY 319 LysLeuLeuAlaAlaH1sCYsPheH1sG1nH1sG1nTyrLeuMetProG1yAsnValVal 338
 Db 1242 AAGCTCTCTGGCAGCCCATGCTTCAATCCAGACCACTACCTCATGCTGAGATGTGTG 1301
 QY 339 GluGluLeuThrLysAlaMetAspLeuH1sAlaValLeuAlaGluLeuProArgProArg 358
 Db 1302 GAGGAGAGAACCAAGGCGCATGACCTGCACGGGCTTGTGGCGAGCTTCCCGGCCCTCC 1361
 QY 359 ArgProProLeuG1nTrrArgTrrYsSerG1yValAsrPheMetG1yLeuAspLysPheLeu 378
 Db 1362 AGGCCCCCGCTGCAGTGGCGGTACTGTGCAGTCTCTTATGAGCTGACAAATTCCTG 1421
 QY 379 GluAspValArgAsnG1y1leTrrYsProLeuMetAsnPheAlaAlaThrArgProLeuG1y 398
 Db 1422 GAGGATGTCAAGATGGAATCTAACCACTGATGAATCTTGCAGGCCACTGCACCCCTGGGG 1481
 QY 399 LeuProArgValLeuAlaProProProG1yGluValGluLysAlaLysThrProThrPro 418
 Db 1482 CTGCCCCGTGTCTGCTGCCCCCAGGAGGTCCAAAGGCCCAAGCCCGGACGCA 1541
 QY 419 GluProPheAspSerG1yThrArgLysVal1leG1yMetG1yCYsAsnLeuG1yArgSer 438
 Db 1542 GAGCCCTTGAATCTGAGACCAAGAAAGTCAATCCAGATCAAGTGCACCTGAGAGAAC 1601
 QY 439 GluAspLysAlaArgTrrH1sLeuThrLeuLeuLeuValLeuG1yAspArgLysH1sArg 458
 Db 1602 GAGGACAAAGCCCGCTGCATCTCACTGTCTGTGCTGCTGGAAGACCGGCTGCACCGG 1661
 QY 459 GluLeuThrTrrYsAspLeuLeuProThrAspSerAlaGluAspLeuAlaSerGluLeuVal 478
 Db 1662 CAGCTGACCTTCACTGCTCCCAAGACGACGCGCCCAAGACCTGCTCGAGCTCGTG 1721
 QY 479 H1sTrrG1yPheLeuH1sG1y----- 485
 Db 1722 CACTATGCTTCTTCCACGA-AGTGGCTGGGCGGTGGCGCGGCTTGGCGGAGGAGGCGC 1780
 QY 486 -----AspAspArg 488
 Db 1781 GCACGGGGGACGGCGCCCTCCGTCGCCCATGCTCCCTCTTCCGACGAGACGACCGG 1840

QY 489 MetLysLeuAlaAlaPheLeuG1ySerThrPheLeuLysTrrArgG1yThrGluAla 507
 Db 1841 ATGAAGCTGGCCGCTTCTTGGAGAGACCTTCTCAAGTACAGTGGAGCCAGGCC 1897
 RESULT 7
 ID ADV97751
 XX ADV97751 standard; cDNA; 3147 BP.
 XX
 AC ADV97751;
 XX
 DT 24-MAR-2005 (first entry)
 XX
 DE cDNA sequence encoding a murine protein kinase Seq 71.
 XX
 KW gene; ser; protein kinase modulator; bioinformatics; diagnostic; cancer;
 KW central nervous system disease; psychotic disorder;
 KW neurological disorder; neurodegenerative disease; metabolic disorder;
 KW cardiovascular disease; inflammatory disorder; gene therapy; cytostatic;
 KW antimigraine; analgesic; endocrine-gen.; nootropic; tranquilizer;
 KW hypotensive; hypertensive; neuroprotective; antiparkinsonian; vitruicide;
 KW fungicide; antibacterial; antidiabetic; anorectic; antiarteriosclerotic;
 KW ophthalmological; antiinflammatory; antiarthritic; antineumatic;
 KW antiasthmatic; osteopathic; antiporiatic; immunosuppressive;
 KW cardiovascular-gen.; vasotropic; antiallergic; gastrointestinal-gen.;
 KW cns-gen.
 XX
 OS Mus sp.
 XX
 PN WO200500200-A2.
 XX
 PD 06-JAN-2005.
 XX
 PF 07-MAY-2004; 2004WO-US014421.
 XX
 PR 09-MAY-2003; 2003US-0469014P.
 XX
 PA (SUGB-) SUGEN INC.
 XX
 PI Caenepeel S, Manning G, Charyczak G, Grigoriev I;
 XX
 DR WPI; 2005-06508/07.
 XX
 DR P-PSDB; ADV97865.
 XX
 PT New isolated, enriched, or purified kinase nucleic acids and
 PT polypeptides, useful for diagnosing or treating, e.g. cancers,
 PT neurological and neurodegenerative diseases, cardiovascular disease, or
 PT inflammatory disorders.
 XX
 PS Claim 10; SEQ ID NO 71; 300bp; English.
 XX
 CC This invention relates to novel isolated, enriched or purified nucleic
 CC acid molecules that encode kinase polypeptides. Specifically, it refers
 CC to a bioinformatics strategy used to identify mammalian members of the
 CC protein and lipid kinase families. The present invention provides methods
 CC for identifying a substance that modulates the activity of a kinase
 CC polypeptide, as well as a method for the detection of a kinase nucleic
 CC acid in a sample as a diagnostic tool for a disease or disorder.
 CC Furthermore, it describes generation of a knock-out mouse whose genome is
 CC disrupted by recombination at a nucleic acid sequence such that it
 CC produces a phenotype, relative to the wild-type, that exhibits an absence
 CC of kinase activity. The nucleic acids and polypeptides given in the
 CC specification are useful for the diagnosis and treatment of cancer;
 CC central or peripheral nervous system diseases, psychotic and neurological
 CC disorders, neurodegenerative diseases, metabolic and neurological
 CC cardiovascular disease or inflammatory disorders. As such, they can be
 CC used for gene therapy purposes and compositions exhibit cytostatic,
 CC antimigraine, analgesic, endocrine-gen., nootropic, tranquilizer,
 CC hypotensive, hypertensive, neuroprotective, antiparkinsonian, vitruicide,
 CC fungicide, antibacterial, antidiabetic, anorectic, antiarteriosclerotic,
 CC ophthalmological, antiinflammatory, antiarthritic, antineumatic,
 CC antiasthmatic, osteopathic, antiporiatic, immunosuppressive, cardiac-
 CC gen, vasotropic, antiallergic and gastrointestinal-gen. This
 CC polynucleotide is a cDNA sequence that encodes a murine protein kinase of

XX P1 P1owman GD, Martinez R, Whyte D, Sudersanam S;
XX DR MPI: 2001-032161/04.
XX P-PSDB; AAB65654.
XX PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
XX PT treating immune-related diseases and disorders, cardiovascular disease,
XX PT neurodegenerative diseases and/or cancers.
XX PS Disclosure, Fig 2, 310pp, English.
XX CC The present sequence encodes a novel protein kinase. The nucleic acids
XX CC and the protein kinases they encode may be used in the treatment and
XX CC diagnosis of diseases associated with inappropriate kinase expression
XX CC such as immune-related diseases and disorders, cardiovascular disease,
XX CC neurodegenerative diseases and/or cancers. The nucleic acids and
XX CC complementary sequences may also be used as DNA probes in diagnostic
XX CC assays. The kinase polypeptides may be used as antigens in the production
XX CC of antibodies of kinase expression and activity. Anti-kinase antibodies
XX CC and kinase antagonists may also be used to down regulate kinase
XX CC expression and activity. Diseases related to kinase expression and
XX CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
XX CC disorders, complications of organ transplantation, myocardial infarction,
XX CC immune disorders, cardiomyopathies, strokes, renal failure, oxidative-
XX CC stress related disorders, chronic inflammatory bowel disease, chronic
XX CC inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,
XX CC psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive
XX CC disorders
SQ Sequence 3304 BP; 736 A; 1019 C; 841 G; 708 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 8.55e-233 Length: 3304
Score: 2409.00 Matches: 460
Percent Similarity: 98.72% Conservative: 2
Best Local Similarity: 98.29% Mismatches: 0
Query March: 90.22% Indels: 6
DB: 4 Gaps: 1
US-10-618-941-99 (1-507) x AAF44681 (1-3304)
QY 40 ATGATGUGLUNVALAaNGINGLYaNMecProGLYLeuGInSerThrPheLeuAlaMet 59
DB 1 CAGAGAGAGAGAGTAAACCAAGGAACATCCAGGGCTTCAAGACCTCTTACGCCATG 60
QY 60 AepThrgUGLUGLYValGluValATrPaNGInLeuNHISpHeGlyAspArgLYaLa 79
DB 61 GACACGGAGAGAGGGGTAGAGGTGTGTGGAACGAGCTCCTTGGAGACAGAAAGCC 120
QY 80 PheAlaAlaNHieGLUGLYuLYaIleGInThrValPheGLInLeuValIleuNHIS 99
DB 121 TTCGGGGGCGACAGAGAGAAAGATCCAGACCGGTTCGAGCACTGGTGTGACACAC 180
QY 100 ProaNIleValLYeLeuNHISerLYrTrPLeuAaPThrSerGlyAlaCyBaLaArgVal 119
DB 181 CCGAACATCGTAAGTTCACAAAGTACGTGGATACCTCTGAGGGCCGCGAGGGTTC 240
QY 120 IlePheIleThrGluTYrValSerSerGlySerLeuLYeGLInPheLeuLYeThrLYs 139
DB 241 ATCTTCATCAGACGATGATGTGCATCAGGCACTTCAGCAATTCCTCAAAAGACCAAG 300
QY 140 LYaBaNHISLYaLIMecAaNHIAArgAlaTrPLYaArgTrPYrThrGInIleuSer 159
DB 301 AAGAACCACAAAGGCATGAACCCCGGGGCTGGAAGCGGTGTGACGAGATCTGTCT 360
QY 160 AlaLeuSerPheLeuNHISaLaCySeSerProProlIleIleHISGLYAsnLeuThrSerA 179
DB 361 GCGCTCAGCTTCTCTGACGCTGACGCCCCCAATCATTCACAGGAACCTGACCGAGAC 420
QY 180 ThrIlePheIleGInHISaNHISaNHISLeuLYeLYaIleGlySerValTrPHeArgIlePhe 199
DB 421 ACCATCTTCATTCACACACAGGGCTCATCAAGATCGGCTCGGTGTGACCGCAATCTTC 480

QY 200 SerAaNHISaLaLeuArgProProThrAlaLeuProAaPheAspLeuArgSerProIleArgAla 219
DB 481 TCCAAAT-----GCATCTCAAGATGATCTCGAAGCCCATCCGCGCT 522
QY 220 GluArgGUGLUGLYaNHISpHePheProProGLYTYrGlyValAlaAaP 239
DB 523 GAGCGAGAGAACTTCGAAACCTGCATCTTCCCCAGAGTATGAGAGGTGGCCGAT 582
QY 240 GlyThrAlaValAspIlePheSerPheGlyMetCYaLaLeuGInuMecAlaValLeuGIn 259
DB 583 GCGACCGCTGTGACATCTTCTCTTGGAGTGTGCGCTGAGATGCTGACTGAA 642
QY 260 IleGInThrAspGlyAspThrArgValATrGUGLUNAlaIleAlaArgAlaArgHISer 279
DB 643 ATCCAGACCAATGGGACACCCGGGTCAAGAGAGGCCATTTGCTCGCCAGGCACTCG 702
QY 280 LeuSerAaPProAaMecArgGluPheIleLeuCYsCYeLeuAlaArgAaPProAlaArg 299
DB 703 CTGAGTGAACCCCAACATGGGGAGTTCATCTTGTCTGCTGGCCCGGAACTTCCCGC 762
QY 300 ArgProSerAlaNHISerLeuLeuPheNHISaArgValIleuPheGluValHISerLeuLYs 319
DB 763 CCGCCCTTGCCCAAGAGCTCTCTTCCACGCGGTGCTTTCGAGGTGCATCGGTGAAG 822
QY 320 LeuLeuAlaAlaNHISpPheIleGInHISGInTYrLeuMecProGluBaNHISaValGlu 339
DB 823 CTCCTGGAGCCCACTGCTTCATCCAGCACCGTATCTCATCTCGAGAAATGTGTGGAG 882
QY 340 GluLYeThrLYaLIMecAspLeuNHISaLaValLeuAlaGluLeuProArgProArg 359
DB 883 GAGAGACCAAGGACATGACCTGACGGGGCTTGGAGAACTTCCCGGCGCCGACAG 942
QY 360 ProProLeuGInTrPArgTYrSerGlyValSerPheMetGluLeuAspLYePheLeuGIn 379
DB 943 CCCCCGCTGACATGGCGGATCTCGAAGTCTCTTCAAGAGCTGACAAATTCCTGGAG 1002
QY 380 AspValaLgaNHISLYrTYrProLeuMecAaPheAlaAlaThrArgProLeuGlyLeu 399
DB 1003 GATGTCAAGAAATGAATCTACCCACTGATGAATTTGACAGCCATCCCTGGGGCTG 1062
QY 400 ProArgValIleuAlaProProProGluGluValGInLYaLaLYeThrProThrProGlu 419
DB 1063 CCCGTGTGTGGCCCCACCCCCGAGAGGTTCAAAGGCCCAAGACCCCGACGACAG 1122
QY 420 ProPheAaPSerGluThrArgLYeValIleGInMecGInCYaAsnLeuGluArgSerGlu 439
DB 1123 CCTTGACTCGAAGACAGAAAGTCAATCCAGATGCAGTGCACTCGAGAGAAAGCGAG 1182
QY 440 AspLYaLaArgTrPHeIleuThrLeuLeuLeuValIleuGluAaPArgLYeNHISaArgGIn 459
DB 1183 GACAAAGGGCGGTGCATCTCACTGTGTGTGTGTAAGAACGGGTGTGACCCGGCAG 1242
QY 460 LeuThrTYrAaPLeuLeuProThrAaPSerIleGInaPLeuAlaSerGluLeuValHIS 479
DB 1243 CTGACCTAGACCTGTCTCCAAAGGACAGCGCCGAGACCTGCTCGAGGTCTGTGAC 1302
QY 480 TYrGlyPheLeuNHISaLaAspAaPArgMecLYeLYaAlaIlePheLeuGInSerThrPhe 499
DB 1303 TATGGCTTCTCTCAAGAGAGACGACCGGATGAAGTGGCGGCTTCTGAGAGACCTTC 1362
QY 500 LeuLYeTYrArgGlyThrGInAla 507
DB 1363 CTCAAGTACCGTGGGACCCAGGCGC 1386
RESULT 9
AD129379
ID AD129379 standard; cDNA, 3304 BP.
XX AD129379;
AC XX
DT 22-APR-2004 (first entry)
XX

DE Human MARK3-associated cDNA #45.
XX Human; ss; antisense gene therapy; MARK3;
KM MAP/microtubule affinity-regulating kinase 3; cancer;
KW Alzheimer's disease; neurodegenerative disorder;
XX hyperproliferative disorder; cytostatic.
OS Homo sapiens.
XX US2003232771-A1.
XX
XX
XX 18-DEC-2003.
XX
XX 17-JUN-2002; 2002US-00174319.
XX
XX 17-JUN-2002; 2002US-00174319.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Ward DT, Freiler SM, Dobie KW;
XX WPI; 2004-052189/05.
XX P-PSDB; ADI29261.
XX
XX
XX New antisense compound targeted to a nucleic acid molecule encoding
PT microtubule-affinity-regulating kinases (MARK3), useful for modulating
PT expression of MARK3 or for treating cancer or Alzheimer's disease.
XX
XX Disclosure; Fig 2; 233pp; English.
XX
XX The invention relates to a compound comprising a sequence comprising 8-80
XX base pairs (bp) targeted to a nucleic acid encoding MARK3
XX (MAP/microtubule affinity-regulating kinase 3), that specifically
XX hybridizes with the nucleic acid encoding MARK3 and inhibits expression
XX of MARK3, i.e., is an antisense oligonucleotide (AO). Also included are a
XX composition comprising the compound and a carrier or diluent, inhibiting
XX the expression of MARK3 in cells or tissues, treating an animal having or
XX suspected of having a disease or condition associated with MARK3 and
XX screening for an antisense compound. The antisense oligonucleotide is
XX useful for preparing a composition for treating hyperproliferative
XX disorder, particularly cancer and neurodegenerative diseases e.g.,
XX Alzheimer's disease. The present sequence is a MARK3 associated cDNA
XX included in the figures but not mentioned anywhere else in the
XX specification.
XX
XX Sequence 3304 BP; 736 A; 1019 C; 841 G; 708 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 8,556-233 Length: 3304
Score: 2409.00 Matches: 460
Percent Similarity: 98.72% Conservative: 2
Best Local Similarity: 98.29% Mismatches: 0
Query Match: 90.22% Indels: 6
DB: 12 Gaps: 1
US-10-618-941-99 (1-507) x ADI29379 (1-3304)
QY 40 Argatggtguglunvalasnglmglyasnmecploglyleuglnsethrphelenu1anet 59
DB 1 CAGAGAGAGAAAGGTAAACCAAGGAAACATGCGAGGGCTTCAAGACACCTTCTGAGCCATG 60
QY 60 Asprrhrglugluglyvalgluvalvalttrpaenglulenuhlsphneglyasprgllyala 79
DB 61 GACACGAGAGAGGGGGGTAGAGGTGTGTGTGAACGAGCTCCACTTCGAGACAGGAAGGCC 120
QY 80 Phealaalanhlsagluluyelleglnthrvallphegluglnleuvalleuvalasphls 99
DB 121 TTCGGGGGCGACGAGAGAAAGATCCAGACCGGTTCGAGCACTGTCTGTGTGACAC 180
QY 100 ProaenillevalylsleuuhlslyatrrtleuaprrhsersglualaCyaalaaargval 119
DB 181 CCGAACATCTGTAAGTTCACAAAGTACTGGCTGATACCTCTGAGGCGCTGGCGAGGGTTC 240

QY 120 lIepheillethrgruglurvalserseglYserleuylsghlnpheleuylvalethrlyls 139
DB 241 ATCTTCATCAGAGTACGTCTCATCAGCAGCCCTCAAGCAATTCCTCAAAAGACCAAG 300
QY 140 lysasnhslysalametasnalaaargalatrplysartgtrpCysrthnglnlleuser 159
DB 301 AAGAACCAAGAGGCAATGAAGCCCGGGCCCTGGAAGCGCTGTGTCAACGACATCTGTCT 360
QY 160 AlauserpheleuuhlsalaCysSerProbroliellenhlsglyamleuthserasp 179
DB 361 GCGCTCAGCTTCGTCAGCGCTCGAGCCCCCAATCATCCAGGAACTTCAACAGCGAC 420
QY 180 ThrllpheelleglnhsaenglyleuullleuylleglYserValttrphlsargllphe 199
DB 421 ACCATCTTATTCAGCAACAGGCTCATCAAGATCGGCTCGGTGACCAATCTTC 480
QY 200 SerasnalaleuargProthrothralaleuProasphleuargserProllaarghla 219
DB 481 TCCAAAT-----GCACTTCAGATGATCTCGAAGCCCATCCGCGCT 522
QY 220 GluargglugluLeuargamleuuhlsPhePheProloglurlyglYgluvala1aap 239
DB 523 GAGCGAGAGAACTTCGGAACCTGCACCTTCTCCCGCAGATGAGAGGTGGCGCAT 582
QY 240 GlythrAlaValaapllPheSerPheglYmetCysAlaleuGlumetAlaValleuglu 259
DB 583 GGGACCGGTGGACATCTTCTTTGGATGTGTGGCTGGAGAGGTGCTGTGAA 642
QY 260 lIeglnthraenglyasprtharvalthrgluglu1a1lea1aargalaaarghseser 279
DB 643 ATCCAGACCAATGGGACACCCGGGTCAAGAGAGGCCAATGCTCGGCGCAAGCACTCG 702
QY 280 LeuserasproasnmecarglupheilleucysCysleu1a1aargapProllaarg 299
DB 703 CTGAGTACCCCAACAGCGGAGTTCATCTTTGCTGCGCGCGGACCTGCGCGC 762
QY 300 ArgproserlanssartleuPhehlsargValleuPhegluvalhlsSerleuyls 319
DB 763 CCGCCCTGTGCGACAGCTCTCTTCCACCGGCTGCTTCGAGGTGCATCTGCTGAAG 822
QY 320 leuenu1a1anhlsCysPheilleglnhlsaglntYrleumecProgluamvala1glu 339
DB 823 CTCCTGGACAGCCACTGCTTCATCAGACCAAGTAACCTCACTGAGAAATGTGTGAG 882
QY 340 GlulYthrlysalametasprleuuhlsAlaValleu1a1gluleuProargProargarg 359
DB 883 GAGAAAGACCAAGGCCATGGAACCTGCACGCGGTCTTGCGGAGCTTCCCGGCGCGCAG 942
QY 360 ProProleuglntrpargtyrrsergluvalserPheMetglulenuaplyasphleuglu 379
DB 943 CCCCCGCTGCAGTGGCGGTACTCGGAAGTCTCTTCATGAGGTGACAAATTCCTTGAG 1002
QY 380 AsprValarganglylletYrProleumecAsnphela1a1atrpaargProleuglyleu 399
DB 1003 GATGTACGAATGGAATCTAACCTGATGAACCTTTCAGAGCACTGCACCCCTGGGGCTG 1062
QY 400 ProargValleu1a1aProProProglugluvalgluylsAla1a1YthrProthProglu 419
DB 1063 CCCCCGTGTGCGCCCAACCCCGAGAGAGGTCCAAAGGCCCAAGACCCCAACGCAAG 1122
QY 420 ProPheasprserglutrrarglyval1leglmetglnCysaenleuGluaargserglu 439
DB 1123 CCGTTGACTCTGAGACCAAGAAAGTATCCAGATGCAAGTGCACCTGGAAGAAAGCGAG 1182
QY 440 Asprlysa1aargtrphlsleuThrleuLeu1a1leuGluaapargleuuhlsargln 459
DB 1183 GACAAAGGCGGCTGCATCTCATCTGCTGTGTGCGAAGACCGGCTGCACCGGAG 1242
QY 460 LeuThrTyraPheleu1a1aProthraasprer1a1aasprleu1a1asergluLeuvalhls 479
DB 1243 CTGACCTAACACCTGCTCCCAACAGGACGCGCCAGGACCTCGCTCGAGAGCTGTGAC 1302
QY 480 TyrglyPheleuuhlsglu1aasprargmetlyleu1a1a1aPheleuglnsetThrPhe 499


```

Db      1303 TATGCTTCTCTCCACGAGGACGCGATGAGCTGCCGCTTCTTGAGAGCACTTC 1362
QY      500 LeuLYTyrArgGLYThrGlnIa 507
Db      1363 CTCAGTACCGTGGAGCCAGGCC 1386

RESULT 10
ADV35041
ID      ADV35041 standard; cDNA; 3538 BP.
AC      ADV35041;
XX
XX      10-FEB-2005 (first entry)
DT
XX
DE      cDNA differentially expressed in the presence of valproate SeqId117.
XX
XX      valproate; 89; multi-parameter high throughput screening; MPHTS;
KW      disease signature; neuropsychiatric; neurodegenerative; schizophrenia;
KW      bipolar affective disorder; BAD; autism; Parkinson's;
KW      Alzheimer's disease; neuroleptic; nootropic; antiepileptic; antidepressant.
XX
XX      Unidentified.
OS
XX      US2003096264-A1.
PN
XX
XX      22-MAY-2003.
PD
XX
XX      18-JUN-2002; 2002US-00175523.
PF
XX
XX      18-JUN-2001; 2001US-0299151P.
PR      07-SEP-2001; 2001US-0317828P.
PR      25-SEP-2001; 2001US-0325150P.
PR      14-NOV-2001; 2001US-0335047P.
PR      18-JAN-2002; 2002US-034936P.
PR      04-MAR-2002; 2002US-0361834P.
XX
XX      (PSTC-) PSYCHIATRIC GENOMICS INC.
PA
XX
XX      Alter CA, Brockman JA, Evans D, Hook D, Klimczak LJ, Laeng P;
PI      Palfreyman M, Rajan P;
XX
XX      MPI; 2004-118903/12.
DR
XX
XX      Identifying a compound that can treat disease or disorders, such as, a
PT      neuropsychiatric disorder e.g., schizophrenia, or autism, comprises
PT      determining the expression of one or more efficacy genes in a cell
PI      contacted with the test compound.
XX
XX      Claim 9; SEQ ID NO 117; 39pp; English.
PS
XX
XX      This invention relates to a novel screening method identified as a multi-
CC      parameter high throughput screening (MPHTS) assay. Specifically, it
CC      refers to an assay that utilizes the disease signature of a plurality of
CC      specific genes associated with a particular disease, and identifies
CC      differential expression between those cells taken from individuals
CC      affected by that disease and those that are not affected. The present
CC      invention then describes the screening of candidate pharmaceutical
CC      compounds to identify those that have a potential therapeutic benefit for
CC      the treatment of neuropsychiatric and neurodegenerative disorders
CC      including schizophrenia, bipolar affective disorder (BAD) and autism, as
CC      well as Parkinson's and Alzheimer's disease. Accordingly, the compounds
CC      of this invention exhibit various activities including neuroleptic,
CC      nootropic, antiepileptic and antidepressant. Furthermore, the screening
CC      method used in MPHTS will be automated, such that a large number of test
CC      compounds may be rapidly screened with a minimal amount of labour and
CC      effort. This polynucleotide is the cDNA sequence of a gene that is
CC      differentially expressed in the presence of the therapeutic compound
CC      valproate, given in an exemplification of the invention.
CC
SQ      Sequence 3538 BP; 806 A; 1076 C; 928 G; 728 T; 0 U; 0 Other;
Alignment Scores:

```

```

Pred. No.: 1,33e-228
Score: 2368.00
Percent Similarity: 87.17%
Best Local Similarity: 86.79%
Query Match: 88.69%
DB: 13
Gaps: 2

US-10-618-941-99 (1-507) x ADV35041 (1-3538)
QY      40 ArgArgGluGlnValAsnGlnIysAsnMetProGlyLeuGlnSerThrPheLeuAlaMet 59
Db      1 CAGAGAGAGAAAGTAAACCAAGGAAACATGCGAGGCTTCAGAGCACTTCCTAGCCATG 60
QY      60 AspThrGluGluGlyValGluValValTTPaAsnGluLeuHisPheGlyAspArgGlyAla 79
Db      61 GACACGAGAGAGAGGGGTAGAGGTGTGTGTGAACGAGCTCCACTTCGAGACAGAGAGGCC 120
QY      80 PheAlaAlaHisGluGluValValIleGlnThrValPheGluGlnLeuValLeuValAspHis 99
Db      121 TTCGCGCGGACAGAGAGAGAGATCCAGACCGTGTTCGAGCAGCTGTGTGTGTGTGTGTGT 180
QY      100 ProAsnIleValIysLeuHisIlySerTyrTrpLeuAspThrSerGluAlaCysAlaArgVal 119
Db      181 CCGAACATCTGTAAGTTGACAAAGTACTGGCTGATACCTCTGAGAGCCCTGGCGAGAGGCTC 240
QY      120 IlePheIleThrGluTyrValIserSerGlySerLeuLysGlnPheLeuLysThrIlys 139
Db      241 ATCTTCATCATCAGAGTACGTGTCTATCAGGACCTCAAGCAATTCCTCAAAAAGACAAAG 300
QY      140 LysAsnHisIlyValAlaMetAspAlaArg----- 148
Db      301 AAGAACCAACAGGCGCATGACGCCGGGTATGGGAGAGCGGGCTGGGAGCCACGAGGGAGC 360
QY      148 ----- 148
Db      361 AGAGAGGGGTGGGCGACGCTCGGGGACTCGGAGTGTGAGGGGGGTGCCCGGCGGCTCGG 420
QY      148 ----- 148
Db      421 ACAGAGGCTGGCGAGATGCGGGGCGGCTCCGAGGCCACGCCCTCTCTCGGCGCC 480
QY      149 ----- 148
Db      481 ACCGACCGAGAGTCTGCGCTCCGCCGCGAGCGCTGAGCGCTGGTGCACGACGATC 540
QY      158 LeuSerAlaLeuSerPheLeuHisIlyCysSerProIleIleHisGlyAsnLeuThr 177
Db      541 CTGTCTGGGCTCAGCTTCTCTGACAGCGCTGACGCCCCCAATCATTCACGAGAACCTGACC 600
QY      178 SerAspThrIlePheIleGlnHisAsnGlyLeuIleLysIleGlySerValTrpHisArg 197
Db      601 AGCGACACATCTTATTCAGACAAACGCGCTCATCAAGATGAGGTGCGGTGGACCGA 660
QY      198 IlePheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIle 217
Db      661 ATCTTCTCAAT-----GACTCTCAAGATGATCTCGAAAGCCCATC 702
QY      218 ArgAlaGluArgGluGluLeuArgAsnLeuHisPhePheProProGlyTyrGlyGluVal 237
Db      703 CGCGCTGACGAGAGAACTTCGAACTTCGCTTCTCCGCCAGATATGAGAGAGGTG 762
QY      238 AlaAspGlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaVal 257
Db      763 GCCGATGGAGACCGCTGTGATCTTCTCTTGGGATGTGTGGCTGAGATGTGCTGTA 822
QY      258 LeuGluIleGlnThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArg 277
Db      823 CTGGAATTCAGAACCAATGAGGAGACCCGCGGTCAACAGAGGAGCCATTGCTCGCCACAG 882
QY      278 HisSerLeuSerAspProAsnMetArgGluPheIleLeuCysValLeuAlaArgAspPro 297
Db      883 CACTGCTGAGTGAACCCAAATGCGAGAGTTTATCTTGTGCTGCTGCGCGGAGACCT 942

```

QY	228	AAAGAGGProSerAlaHisSerLeuLeuPheHisArgValLeuPheGluValHisSer	317
Db	943	GCCCCCGGCGCTCTGCGCCACAGCCCTCTTCCACCGGAGCTCTTGAGAGTGACATCG	1002
QY	318	LeuValLeuLeuAlaAlaHisCysPheIleGlnHisGlnIleTyrLeuMetProGluAsnVal	337
Db	1003	CTGAAGCTCTCTGGCAGCCCACTGCTTACCTCCAGCACCACTACTCATGCTCGAGATGTG	1062
QY	338	ValGluGluValThrIleValMetCysPheLeuHisAlaValLeuAlaGluLeuProArgPro	357
Db	1063	GTGGAGGAGAAACCAAGGCGCATGAGCCCTGCACCGGAGTCTTGCGGAGCTTCCCGGCC	1122
QY	358	ArgAGGProProLeuGlnThrArgTyrSerGluValSerPheMetGluLeuAspIlePhe	377
Db	1123	CGCAGGCCCCCGCTGACGTGGCGGTACTCGAAGTCTCTTCATGAGCTGACAAATTC	1182
QY	378	LeuGluAspValArgAsnGlyIleTyrProLeuMetAsnPheAlaAlaThrArgProLeu	397
Db	1183	CTGGAGGATGTACGAGATGGAAATCTACCACTGATGAACTTTGACGCCACTGACCCCTG	1242
QY	398	GlyLeuProArgValLeuAlaProProProGluGluValGlnValAlaIleValThrProThr	417
Db	1243	GGGCTGCCCCGCTGTGTGGCCCCACCCCGAGAGAGTCCAAAGGCCAAGCCCCGACG	1302
QY	418	ProGluProPheAspSerGluThrArgIleValIleGlnMetGlnCysAsnLeuGluArg	437
Db	1303	CCAGAGCCCTTTGACTCTGAGACCAAGAAAGGTCACTCCAGATGCAACTGCAACTGAGAGA	1362
QY	438	SerGluAspValAlaArgTrpHisLeuThrLeuLeuValLeuGluAspArgLeuHis	457
Db	1363	ACCGAGACAAAGCCGCGCTGGCATCTCACTTCTGAGTCTGAAAGACCGGCTGCAC	1422
QY	458	ArgGlnLeuThrTyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeu	477
Db	1423	CGGCAAGTCACTTACGACCTGTCTCCCAACGACAGCCGCCAGGACCTGGCTCGAGCTC	1482
QY	478	ValHisTyrGlyPheLeuHisGluLeuAspArgMetIleValAlaPheLeuGluSer	497
Db	1483	GTGCACACTATGGCTTCTTCCACAGAGACGCGACGATGAAGCTGGCGCTTCTCTGAGAGC	1542
QY	498	ThrPheLeuValTyrArgGlyThrGlnAla	507
Db	1543	ACCTTCTCTCAAGTACCGTGGAGCCCAAGGCC	1572
RESULT 11			
AEA61757			
ID	AEA61757	standard, cDNA; 3538 BP.	
AC	AEA61757;		
DT	11-AUG-2005	(first entry)	
XX			
DE		Human nuclear receptor binding protein 2 (NRBP2) cDNA.	
XX			
KM		Nuclear receptor binding protein 2; GTPase modulator; cancer; neoplasm,	
KW		cytostatic; gene therapy; antibody therapy; drug screening; diagnosis;	
KW		gene; ss.	
OS		Homo sapiens.	
XX			
XX	Key	Location/Qualifiers	
FT	CDS	799..1575	
FT		/*tag= a	
FT		/gene= "NRBP2"	
FT		/product= "Human nuclear receptor binding protein 2"	
XX			
PN	WO2005052130-A2.		
XX			
PD	09-JUN-2005.		
XX			
XX	23-NOV-2004; 2004WO-US039708.		
XX			
PR	24-NOV-2003; 2003US-0524541P.		

Seq	Sequence	Length	Matches	Conservative	Mismatches	Indels	Gaps
40	ARGAAGGUGLGNVALASNGLNGLYASNMETPROGIVLEGLNLSERTHRHELENAIAME	3538	2368	87.17%	2	68	2
1	CAGAGAGGAAAGGTAAACCCAAAGGAAACAAAGCCAGGCGCTTCAAGACACTTCCATGACATG	460	2368	87.17%	2	68	2
60	ASPTTNGUGLNGLYVALGIVLVALVATTPAENGULLEUHSFPHGLYASPARGLYSALA	79	2368	87.17%	2	68	2
61	GACACGGAGAGAGGGGGGTAGAGGTGGTGTGGAAACGAGCTCCACTTGGAGACAGAAAGGCC	1200	2368	87.17%	2	68	2
80	PHALAAALAHISGLIGLULYSILLEGINTHRVALPHEGLIGLILEUVALILEUVALIAEPHIS	99	2368	87.17%	2	68	2
121	TTCCGGCGCGACGAGAGAGAAAGATCCAGACCGTGTTCGACAGCTGGTGTGGTGGACAC	1800	2368	87.17%	2	68	2
100	PROBANTLLEVALYLYLEUHSILYSYTYTRTLEUASPTHSERGILUALCYSAIAGVAL	119	2368	87.17%	2	68	2
181	CCGAAACATCGTGAAGTTCGACAAAGTACCTGGCTGGATCCTTCGAGGCTTCGCGAGGGTC	2400	2368	87.17%	2	68	2
120	ILEPHEIETHNGULTYRVAISERSEGLYSERILEUYSGLIPHEUYSILYSYTHIYS	139	2368	87.17%	2	68	2
241	ATCTTCATCAGAGAGTACGTGCATCAGGACGCTTCAGCAATTCCTCAAAAAGACCAAG	3000	2368	87.17%	2	68	2
140	LYSAASHISLYSALAMECAASHIAATG	148	2368	87.17%	2	68	2

[illegible]

Db	1363	AGCGAGACAAAGGCGCGCTGGCATCTGCATCTGCTTGGTGGCTGGAGAGACCGGCTGGAC	1422
Qy	458	ArgGlnLeuTrpTyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeu	477
Db	1423	CGGACGCTGACCTTACGACTCTCTCCACAGGACGCGCCCAAGGACTTCGCTCGAGACTC	1482
Qy	478	ValHisTyrGlyPheLeuHisGluAspAspArgMetLysLeuAlaAlaPheLeuGluSer	497
Db	1483	GTGCACTATGGCTTCTCTCCACGAGAGACGACCGGATGTAAGCTGGCGGCTTCTCGAGAGAC	1542
Qy	498	ThrPheLeuLysTyrPArgGlyThrGlnAla	507
Db	1543	ACCTTCTCAAGTACCGTGGAGCCAGGCC	1572
RESULT 12			
ID	AAKS1824	standard, cDNA, 1501 BP.	
XX	AAKS1824;		
AC			
DT	06-NOV-2001	(first entry)	
XX			
DE	Human polynucleotide SEQ ID NO 369.		
XX			
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;		
KM	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;		
KW	tissue growth factor; immunomodulatory; cancer; leukemia;		
KM	nerve system disorder; arthritis; inflammation; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200157190-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	05-FEB-2001; 2001WO-US004098.		
XX			
PR	03-FEB-2000; 2000US-00496914.		
PR	27-APR-2000; 2000US-00560875.		
PR	20-JUN-2000; 2000US-00598075.		
PR	19-JUL-2000; 2000US-00620325.		
PR	01-SEP-2000; 2000US-00654936.		
PR	15-SEP-2000; 2000US-00663561.		
PR	20-OCT-2000; 2000US-00693325.		
PR	30-NOV-2000; 2000US-00728422.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Tang YF, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;		
PI	Ma Y, Zhao QJ, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;		
PI	Xue AJ, Yang Y, Wejhtman T, Goodrich R;		
XX			
DR	WPI; 2001-476283/51.		
DR	P-PBDJ; AAM78691.		
XX			
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful		
XX	in diagnosis and gene therapy.		
XX			
PS	Claim 1; Page 1404-1406; 6221pp; English.		
XX			
CC	The invention relates to polynucleotides (AAKS1456-AAKS3435) and the		
CC	encoded polypeptides (AAM78333-AAM80302) that exhibit activity elating to		
CC	cytokine, cell proliferation or cell differentiation or which may induce		
CC	production of other cytokines in other cell populations. The		
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or		
CC	peptide therapy. The polypeptides have various cytokine-like activities, e.g.		
CC	stem cell growth factor activity, haematopoiesis regulating		
CC	activity, tissue growth factor activity, immunomodulatory activity and		
CC	activin/inhibin activity and may be useful in the diagnosis and/or		
CC	treatment of cancer, leukemia, nervous system disorders, arthritis and		
CC	inflammation. Note: Records for SEQ ID NO 2110 (AAKS2581), 2111		
CC	(AAKS2582) and 3666 (AAM80020) are omitted as the relevant pages from the		
CC	sequence listing were missing at the time of publication		

XX Sequence 1501 BP; 300 A; 504 C; 432 G; 264 T; 0 U; 1 Other;
 SO Alignment Scores:
 Pred. No.: 2,976-212 Length: 1501
 Score: 2201.00 Matches: 421
 Percent Similarity: 98.37% Conservative: 2
 Best Local Similarity: 97.91% Mismatch: 1
 Query Match: 82.43% Indels: 6
 DB: 4 Gaps: 1
 US-10-618-941-99 (1-507) x AAK51824 (1-1501)
 QY 78 LysAlaPheAlaAlaHisGluGluValIleGlnThrValPheGlnLeuValLeuVal 97
 Db ::::|||||
 QY 55 CGTGGCTTGGCGGCGCAGAGAGAAATCCAGACCGTGTGAGACACTGGTGTGGTG 114
 Db GACCAACCGGAACATGTAAGTGCACAAATGCTGGCTGATACCTCGAGGCGTGGCG 174
 QY 98 AspHisProAsnIleValIleLeuHisIleValTyrTrpLeuAspThrSerGluAlaCysAla 117
 Db 115 GACCAACCGGAACATGTAAGTGCACAAATGCTGGCTGATACCTCGAGGCGTGGCG 174
 QY 118 ArgValIlePheIleThrGluTyrValIleSerSerGlySerLeuValPheIleValIle 137
 Db 175 AGGGTCATCTTCATCAGAGATACGTGTATCATGAGGACCTCAAGCAATTCCTCAAAAG 234
 QY 138 ThrIleValSerAsnHisIleValIleMetAsnAlaArgAlaTyrValIleTyrCysThrGlnIle 157
 Db 235 ACCAAGAAACCAACAGGCCATGAAACCGCGGCTGAAAGCGTGGTGCACGAGATC 294
 QY 158 LeuSerAlaLeuSerPheLeuHisAlaCysSerProPheIleIleHisGluValLeuThr 177
 Db 295 CTGTCTGGCTCAGCTTCTGCAACCGCTGCACCCCAATCATCCAGGGAACCTGACC 354
 QY 178 SerAspThrIlePheIleGlnHisIleGlnGlyLeuIleValIleGlySerValIleTyrHisArg 197
 Db 355 AGCAGACACATCTTCATCAGAGATACGTGTATCATGAGGACCTCAAGCAATTCCTCAAAAG 414
 QY 198 IlePheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIle 217
 Db 415 ATCTTCTCAAT-----GCACTTCCAGAGATCTCCGAAAGCCCAATC 456
 QY 218 ArgAlaGluArgGluGluLeuArgValIlePhePheProProGluTyrGluValIle 237
 Db 457 CGCGCTGACGAGGAAGAACTTCGGAACCTTCTTCCCAAGATATGAGAGGTG 516
 QY 238 AlaAspGlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaValIle 257
 Db 517 GCCGATGGAGACCGCTGTGACATCTTCTTGGGAGTGTGCGCTGAGATGCTGTGA 576
 QY 258 LeuGluIleGlnThrAsnGlyAspThrArgValIleThrGluGluAlaIleAlaArgAlaArg 277
 Db 577 CTGGAAATCCAGACCAATGGGAGCAACCGGGTCAAGAGAGGCCATTTGCTGGCCGAG 636
 QY 278 HisSerLeuSerAspProAsnMetArgGluPheIleLeuCysCysValLeuAlaArgAspPro 297
 Db 637 CACTCGCTGAGTGAACCCCAACATGCGGAGATTCATCTTCTGCTGCTGGCCGAGACCT 696
 QY 298 AlaArgArgProSerAlaHisSerLeuLeuPheHisArgValIlePheGluValHisSer 317
 Db 697 GCCCGCCGCGCTTGGCCACAGCTCTCTTCCACCCCGTGTCTTGAAGTGCACGCG 756
 QY 318 LeuValLeuLeuAlaAlaHisCysPheIleGlnHisGlnTyrTrpLeuMetProGluAsnValIle 337
 Db 757 CTGAAGCTCTTGGCGCCCACTGCTTCATCCAGACCGAGTACCTCATCCGAGATGTG 816
 QY 338 ValGluGluValThrValAlaMetAspLeuHisAlaValIleLeuAlaGluLeuProArgPro 357
 Db 817 GTGAGAGAGAACCAAGCCATGCACTTCGACGCGGTCTTGGCGGAACTTCCCGGCGCC 876
 QY 358 ArgArgProProLeuGlnTyrArgTyrSerGluValIleSerPheMetGluLeuAspIlePhe 377
 Db 877 CGCAGGCGCCGCTGCAATGCGGTACTCGAAAGTCTCTTCATGAGACTGACAAATTC 936

QY 378 LeuGluAspValIleArgAsnGlyIleTyrProLeuMetAsnPheAlaIleThrArgProLeu 397
 Db 937 CTGAGAGATGTCAAGAAATGGAATATACCACTGATGAATCTTGACGACCTGACCCCTG 996
 QY 398 GlyLeuProArgValIleAlaProProGluGluValGlnValAlaIleTyrThrProThr 417
 Db 997 GGGCTGCGCCGTGTGTGGCCCAACCCCGAGAGAGGTCCAAAAGGCCAAGACCCCGAGCG 1056
 QY 418 ProGluProPheAspSerGluThrArgIleValIleGlnMetGlnCysAsnLeuGluArg 437
 Db 1057 CCAGAGCCCTTTGACTGTGAGACCAAGAAAGTCTATCCAGATGACGTGCACCTGAGAGA 1116
 QY 438 SerGluAspValIleArgTrpHisLeuThrIleLeuValIleGluAspArgLeuHis 457
 Db 1117 AGCAGAGACAAAGCGCGCTGCTCATCTCTCTTGTGTGTGAGAGACCGGCTGCAC 1176
 QY 458 ArgGlnLeuThrTyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeu 477
 Db 1177 CGGCACTGACCTTACGACCTGCTCCAAACGAGACAGCGCCGACCTGCGGAGCTC 1236
 QY 478 ValHisTyrGlyPheLeuHisGluAspAspArgMetIleValAlaIlePheLeuGluSer 497
 Db 1237 GTGCACTATGCTTCTTCCAGACGACCGGATGAAGCTGGCGCTTCTGAGAGAC 1296
 QY 498 ThrPheLeuValTyrArgGlyThrGlnAla 507
 Db 1297 ACCTTCTCAAGTACGTTGGGACCAAGGCC 1326
 RESULT 13
 AAF29899
 ID AAF29899 standard; DNA; 981 BP.
 AC AAF29899;
 DT 04-APR-2001 (first entry)
 XX Human h15993 DNA.
 DE Human h15993 DNA.
 KM Human; protein kinase; cell growth; tumour; cancer; immune; inflammatory;
 KM respiratory; haematological; bone disorder; ds.
 OS Homo sapiens.
 PN W0200100879-A1.
 XX 04-JAN-2001.
 PD 30-JUN-2000; 2000WO-US018291.
 PF 30-JUN-1999; 99US-00345473.
 PR 01-MAY-2000; 2000US-00562480.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA Hodge MR, Meyers R, Williamson M;
 PI WPI; 2001-061977/07.
 DR New protein kinase polypeptides, nucleic acids and anti-kinase
 PT antibodies, useful for diagnosing and treating e.g. cancer, inflammatory,
 PT immune, cardiovascular and bone disorders.
 XX Claim 1; Fig 13; 93pp; English.
 XX The present invention relates to human protein kinase. The proteins are
 CC from new human genes termed h12832, h14138, h14833, h15990, h15993,
 CC h16341 and h2252. The proteins may be used to identify modulators of
 CC their activity. The proteins may also be used to derive products for the
 CC treatment of cellular growth related disorders, malignancies, cancers,
 CC immune, inflammatory, respiratory, haematological and bone-related
 CC disorders
 CC Sequence 981 BP; 209 A; 298 C; 309 G; 162 T; 0 U; 3 Other;

Score: 1534.50 Matches: 305
Percent Similarity: 73.54% Conservative: 59
Best Local Similarity: 61.62% Mismatches: 102
Query Match: 57.47% Indels: 29
DB: 4 Gaps: 6
US-10-618-941-99 (1-507) x AAH14904 (1-2141)

Oy 17 GIUAAGGUAAPGUSGUGUAAPGUSAPRPIELEUGUGUSERPROCYGLUARG 36
Db 220 GAGGAAGAGAGAGAGAGAGATGATGCTGAGATTTTGGAGAGATCCCGCTGGGGGC 279
Oy 37 TRPGIUYASRGUGUGUValaengIngluammetPROGLUYEUGUSERThrphe 56
Db 280 TGGCGAAGAGGCGAGAGAGAGTGAATCAACGAAATGACAGGATTTGACAGTCATAC 339
Oy 57 LeuAlaMetAspThrGluGluGluValaITPRaNGluLeuHAspGluYAP 76
Db 340 CTGGCCATGATACAGAGAGAGGTGAGAGGTGTGTGGAAAGAGGTACAGTTCTGAA 399
Oy 77 ArgIyValaPheAlaIaHAsgluGluYsIeGIInThrValaPheGluGluValaLeu 96
Db 400 CGCAAGAACTACAGCTCAGAGAGAGAAAGGTTCTGCTGCTGTGATATCTGATCA 459
Oy 97 ValaAspHAspProAsnIleValIyAspLeuHAslySTyTTPLeuAspThrSerGluAlaCys 116
Db 460 TTGGAGCATCTTAAACATTTGTTAAGTTTCAACAAATTTGGGCTGACATTAAAGAGAACAG 519
Oy 117 AlaArgValaIlePheIleThrGluTyValaSerSerGlySerLeuYsGluPheLeuYs 136
Db 520 GCCAGGGTCATTTTATACAGAAATACATGTCATCTGGGAGCTGAGCAATTTCTGAG 579
Oy 137 LysThrIyValaAsnHAslyValaMetAsnAlaArgAlaTPRlyAspTyTPCyThrGln 156
Db 580 AAGAACAAAGAGAACACAGACGATGATGAAAGGATGAGACCTTGGTGGCACACAA 639
Oy 157 IleLeuSerAlaLeuSerPheLeuHAslyValaCysSerProProlleIleHAsglYAsnLeu 176
Db 640 ATCTCTTGGCCCTAAGCTAAGCTGCTGCTGCTGAGCCCCCATATATCAGAGAACTGG 699
Oy 177 ThrSerAspThrIlePheIleGIInHAsngIyLeuIleYsIeGlySerValTPRHis 196
Db 700 AACTGTGACACATCTTATCCAGACAGAGACATCAAGATTGGCTGTGTG----- 753
Oy 197 ArgIlePheSerAsnAlaLeuArgProProThralaLeuProAspAspLeuArgSerPro 216
Db 754 -----GCTCCGACACCATCAACATCAT 777
Oy 217 IleArgAlaGluArgGluLeuArgAsnLeuHAspPheProProGluTyGlu 236
Db 778 GTGAAGACTTGTGAGAGAGAGAGAGATCTACCTTTGACACAGAGATGAGAGAA 837
Oy 237 ValAlaAsp---GlyThralaValaAspIlePheSerPheGlyMetCysAlaLeuGluMet 255
Db 838 GTCACTATGTGACAAAGAGAGAGTGAATCTACTCTTGGCATGTGTGCACTGAGAGTG 897
Oy 256 AlaValaLeuGluIleGIInThrAsngIyAspThrArg--ValThgluGluAlaIleAla 274
Db 898 GCAGGTGCTGAGATTCAGAGGACATGAGATCTCATATGTGGCCACAGAGAGCCATCAGC 957
Oy 275 ArgAlaArgHAspSerLeuSerAspProAsnMetCysGluPheIleLeuCyCysLeuAla 294
Db 958 AGTGCATCCAGCTTCTAGAGAGAGAGAGATCTACAGAGAGAGTTCATCAAGAGCTGAG 1017
Oy 295 ArgAspProAlaArgArgProSerAlaHAspSerLeuLeuPheHAsArgValaLeuPheGlu 314
Db 1018 TCTGAGCTGTGCTCGAGACCAAGAGAGAGAACTTCTGTTCCACCCAGCATTTGTTGAA 1077
Oy 315 ValHAspSerLeuYsLeuLeuAlaIaHAsCyPheIleGIInHAsglTyTleuMetPro 334
Db 1078 GTGCCCTGTCTCAAACTCTTGGCGGCCCATGTGATTTGGAGACCAACACATGATCCCA 1137
Oy 335 GluAsnValaIleGluIyLysThryValaMetAspLeuHAslyValaValaLeuAlaGluLeu 354

Db 1138 GAGAAAGCTGTAGAGAGATACCAAAACATGATTAAGTCCGCTATGCTGAAATTC 1197
Oy 355 ProArg--ProArgArgProProLeuGluITPRArgTySerGluValaSerPheMetGlu 373
Db 1198 CCTGACAGACCAAGAGAGAGAGAGTTCAGACTTTGTACTCTCAAGTACAGAGCTTGAA 1257
Oy 374 LeuAspIyPheLeuGluAspValaArgAsngIyIleTyTPRLeuMetAsnPheAlaIa 393
Db 1258 TTAGATAAATTCCTTGAAGAGTCAAGAAATGGAGATCTATCTCTGACAGCCTTT----- 1311
Oy 394 ThrArgProLeuGlyLeuProArgValaLeuAlaProProProGluGluVala----- 410
Db 1312 -----GGGCTGCTCGGCCCCCAGAGCCACAGCAGAGAGAGATACATCAGCT 1359
Oy 411 -----GInlyValaIySThrProThrProGluProPheAspSerGluThrArgIyAs 427
Db 1360 GTCGGCCCCCTGTGCAAGAGCTCCGACACTTAACAGCTGAGGTGAGATCTGCAAG 1419
Oy 428 ValIleGIInMetGInCyAsnLeuGluArgSerGluAspIyValaArgTPRHisLeuThr 447
Db 1420 GTGGTGTGATGACATGCAACATTAAGTCGGTGAAGAGGAGAGTCAACACACCTGACA 1479
Oy 448 LeuLeuValaLeuGluAspArgLeuHAsArgGluLeuThryTzAspLeuLeuProThr 467
Db 1480 CTCTGCTGAAGTTGAGAGACAACTGAACCGGACCTGAGCTGTGACTGATGCTGCAAT 1539
Oy 468 AspSerAlaGluAspLeuAlaSerGluLeuValaHAsTyGlyPheLeuHAsglYAspAsp 487
Db 1540 GAGAAATATCCCGAGTTGGCGGCTGAGCTGGTGCAGCTGGGCTTCATTATAGAGGCTGAC 1599
Oy 488 ArgMetIyValaAlaIaPheLeuGluSerThrPheLeuYsTyT 502
Db 1600 CAGAGCGGTTGACTCTCTGCTGAAGAGACCTTGAACAAAGTTC 1644

RESULT 15
ABA93734
ID ABA93734 standard; cDNA, 2158 BP.
XX
AC ABA93734;
XX
DT 30-APR-2002 (first entry)
XX
DE Human brain derived cDNA clone fbr2_78d18.
XX
KW Human; foetal brain; foetal kidney; melanoma; testis; amygdala;
XX gene therapy; ss.
XX
OS Homo sapiens.
XX
FN W0200198454-A2.
XX
PD 27-DEC-2001.
XX
PE 25-APR-2001; 2001KO-IB002050.
XX
PR 25-APR-2000; 2000US-0199380P.
XX
PA (GBHU-) GERMAN HUMAN GENOME PROJECT.
XX
PI Wiemann S;
XX
DR WPI; 2002-055860/07.
XX P-PSDB; ABB05696.
XX
PT Human cDNA sequences and clones derived from human fetal brain, fetal
XX kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic
XX screening and therapy.
XX
PS Claim 1; Page 215-216; 611p; English.
XX
CC The present invention describes assemblages and computer readable media
XX comprising novel human cDNA sequences and clones derived from human

CC foetal brain, foetal kidney, melanoma, testis and amygdala cDNA
 CC libraries. AB93702 to AB93766 represent human cDNA sequences from the
 CC present invention which encode the proteins given in AB05662 to
 CC AB05729. The human cDNA sequences and clones can be used in gene
 CC therapy. The clones may be used in a variety of applications, for example
 CC they may be used in profiling assays, for providing large arrays of human
 CC genetic material for implementing large-scale screening strategies and
 CC for treating diseases via gene therapy procedures

XX Sequence 2158 BP; 537 A; 570 C; 576 G; 475 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1-4e-144	Length:	2158
Score:	1534.50	Matches:	305
Percent Similarity:	73.54%	Conservative:	59
Beet Local Similarity:	61.62%	Mismatches:	102
Query Match:	57.47%	Indels:	29
DB:	6	Gaps:	6

US-10-618-941-99 (1-507) x AB93734 (1-2158)

```

OY 17 GIUATGUAAPGUAUSERGUAAPGUAUSERAPPILEUGLUGLUSERPROCYSEGLYARG 36
DB 217 GAGGAAGAAGAAAGTGAAGATGAGTCTGAGATTGGAAGAGTCCCTGTGGGCGC 276
OY 37 TTPGUAUARGAGLUGLUGLUAENGNGLYAARMETPROGLYLEUGLUSERTHRPE 56
DB 277 TGGCAAGAAAGCGAAGAAAGGTGAATCAACGAATGACAGATTAATACAGTGCATC 336
OY 57 LEUAMETAPRTHGUGLUGLUGLYAIGLUAVALTPRANGLUENHISPHLEGUAP 76
DB 337 CTGGCATGATGATACAGAGAGAGTGAAGCTTGTGTGAAATGAGTACAGTCTCTGA 396
OY 77 ARGLYAALAPHEALALAHISGLUGLUGLYLEGINTHVALPHLEGULNEUVALLEU 96
DB 397 CGCAAGAACTACAGACTGCGAGAGGAAAGTTCCGCTGCTGTGATTAATCTGATTC 456
OY 97 VALAPRHSAPRAPHLEVALLEUVALLEUHSLEUVALTPRLEUAPRTHSERGLUACYS 116
DB 457 TTGGAGCATCTTAACCTGTGTAAAGTTTCCAAATATTTGGGCTGACATTAAGAGAAC 516
OY 117 ALAATGVALLEPHELETHRGUUTYRVALSERSEGLYSERLEUENGINPHELEUYS 136
DB 517 GCCAGGCTCATTTTATTCACAAATACATGCTGAGAGTCTGAAGCAATTTCTGAAG 576
OY 137 LYETHLYLYAASHHISLYVALAMETASNAALATGALATPLYARGTTPCYETHRGIN 156
DB 577 AAGACCAAAAAGAACCAACAGATGATGAAGCAATGAGCATGAGCGTTGGTGCACACA 636
OY 157 ILEUASERVALLEUSERPHELEUHSIALCYSERPROFIOLEHISGLYVANLEU 176
DB 637 ATCCCTCTGCGCTTAAGCTACCTGCACTCTGTCAGCCCCCATCATCATCGGAACCTG 696
OY 177 THRSEAPRTHRIIPHELEGINHISANGLYLEUVALLEGLYSERVALTRPHIS 196
DB 697 ACCGTGACACCATCTTCTATCCAGCAACAGGATCATCAAGATTGGCTGTG----- 750
OY 197 ARGILEPHESEASNALEUARGPROPTHRIALEUAPRASPAPLEUARGSERPRO 216
DB 751 -----GCTCCGACACATATCAACAATCAT 774
OY 217 ILEAAGLALGUAUGLUGLUGLEUARGVANLEUHSIPHEPAPRPROGLUYRGLYGLU 236
DB 775 GTGAAGACTTGTGGAAGAGCAAGAAAGATTAACATCTTCTTGACACCAAGATATGAGAA 834
OY 237 VALAIAAPR---GLYTHRALVALAPRIIPHESEPRNEGIMETCYALALEUGLUMET 255
DB 835 GTCACTAATGTGACAAACAGTGAATCTACTCTTTGGCATGTGCTGACCTGAGAGAG 894
OY 256 ALAVALLEUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUG 274
DB 895 GCAAGTGTGAGAGATTCAAGGCAATGAGAGTCTCATATATGTGCACAGAGAACCATCAGC 954

```

```

OY 275 ARGALARGHISERLEUSERAPRAPHMETARGGLUPHELEUCYSECYSLA 294
DB 955 AGTGCATCCAGCTTCTGAAGAGCCCATTAACAGAGGAGTTTATTCAAAAGTGGCGAG 1014
OY 295 ARGAPRPOALARGARGPROSERALAHISERLEUENHISARGVALLEUPHEG 314
DB 1015 TGTGAGCGTGTGCGAGAACCAACAGAGAGAACTCTGTTCACCCAGCATTTGTTGA 1074
OY 315 VALHISERLEUVALLEUVALAHISCYSPHEILEGINHISGLINTYRLEUMETPRO 334
DB 1075 GTGCCCTGCTCAACTCTTCCGCGCCAGCTGATTTGGAGACACAAACATGATCCCA 1134
OY 335 GIUAENVALVALGUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUG 354
DB 1135 GAAACGCTCTAGAGAGATCAACAAACATGATTAAGTCCGCTATGCGCTGAAATC 1194
OY 355 PROARG---PROARGARGPROBLEUGINTPRAGTYRSEGLYVALSERPHEMETGLU 373
DB 1195 CCTGCAGACCAAGAGAGAACCACTTCAAGACTTTGTACTCTCACTGACCACTGAGAA 1254
OY 374 LEUAPRHSAPRAPHLEUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUG 393
DB 1255 TTAGATTAATCTCTGAAGATGTCAGGAATGGATCTATCTCTGACAGCTTT----- 1308
OY 394 THRARGPROBLEUGLYLEUPRARGVALLEUVALPROPRPROGLUGLUGLUGLUGLUG 410
DB 1309 -----GGGCTGCTCTGGCCCAAGCAGCCAGCAGAGAGAGGTGACATCAGCT 1356
OY 411 -----GLNLYALALYETHPRTHPRTHPRTHPRTHPRTHPRTHPRTHPRTHPRTH 427
DB 1357 GTCCGCCCCCTCTGTCAAGCTCCGACCTGAAACCACTGAGAGTGAAGCTGCGAAG 1416
OY 428 VALLEGINMETGLINCYAASHHISGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUG 447
DB 1417 GTGGTGTGAGTGAAGTGAAGCAACTGAGTGGTGAAGAGAGAGTCAACACCACTGACA 1476
OY 448 LEUENLEUVALLEUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUG 467
DB 1477 CTTCGTGAGAGTGAAGAGCAAACTGAACCGGCACTGAGCTGTGACCTGATGCGCAAT 1536
OY 468 AEPSEVALAGLAPRLEUVALSERGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUG 487
DB 1537 GAGAAATATCCCGAGATTGGCGGCTGAGCTGTGAGCTGAGCTGATTAATGAGGCTGAC 1596
OY 488 ARGMETLYLEUVALALAPHELEUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUG 502
DB 1597 CAGAGCCGAGTGTGACTTCTGTGAGAGAGACCTTGAACAAAGTTTC 1641

```

Search completed: January 13, 2006, 21:44:48
 Job time : 779 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 13, 2006, 21:11:56 ; Search time 4609 Seconds
(without alignments)
5146.681 Million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670
Sequence: 1 MAAPAPAPRARRERERED.....RMKLAAPLESTFLKRYGTQA 507

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODBL=frame+pn.model -DEV=xlp
-Q/cgnt2_usfto.spool_p/US10618941/runatc_12012006_070204_14281/app.query.fasta_1.647
-DB=EST -QPMT=fastap -SUFFIX=rc -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdd -LIST=45
-DOCALLGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPM=pro -NORP=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10618941.CCGN.1.1.8010.0@runatc_12012006_070204_14281 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THHEADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gsa1:*
10: gb_gsa2:*
11: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2391	89.6	3163	4	BC068117 Mus muscu
2	2037.5	78.3	3729	4	BC037396 Homo sapi
3	1562.5	56.3	2163	4	AK046142 Mus muscu
4	1537.5	57.6	1608	10	AY408905 Mus muscu
5	1536	57.5	2616	4	HS060563 AL034530 Homo sapi
6	1534.5	57.5	1601	10	AY408904 Pan trogl
7	1534.5	57.5	1608	10	AY408903 Homo sapi

8	1534.5	57.5	2137	4	CR604477	CR604477 full-leng
9	1534.5	57.5	2158	4	HS060150	AL136682 Homo sapi
10	1534.5	57.5	2181	4	CR858659	CR858659 Pongo pyg
11	1520	56.9	2086	4	CR591118	CR591118 full-leng
12	1518.5	56.9	2181	4	AK077440	AK077440 Mus muscu
13	1481	55.5	899	7	CR998804	CR998804 CR998804
14	1403	52.5	927	5	BK410189	BK410189 BX410189
15	1429	46.0	825	3	BI645887	BI645887 60327555
16	1199	44.9	729	6	CD354461	CD354461 UT-M-GMO-
17	1170	43.8	2923	4	AF318376	AF318376 Homo sapi
18	1165.5	43.7	859	7	CK476491	CK476491 AGENCOURT
19	1161	43.5	673	8	DN998004	DN998004 TC107239
20	1160.5	43.5	841	7	CO248961	CO248961 AGENCOURT
21	1126	42.2	716	6	CV105566	CV105566 AGENCOURT
22	1109	41.5	700	7	CN358568	CN358568 170005319
23	1100.5	41.2	1028	1	AL532940	AL532940 AL532940
24	1098	41.1	899	2	BF299723	BF299723 602029226
25	1055.5	39.5	770	7	CN218030	CN218030 RJ4022G03
26	1050.5	39.3	725	2	BG913260	BG913260 602811940
27	1042	39.0	651	6	CB557694	CB557694 AMGNMNC-N
28	1030	38.6	996	3	BI650160	BI650160 603296250
29	1025.5	38.4	911	5	BU508281	BU508281 AGENCOURT
30	1015	38.0	959	5	BQ652622	BQ652622 AGENCOURT
31	1013	37.9	643	7	CN358567	CN358567 170005322
32	1012.5	37.9	1382	8	DN656596	DN656596 CEC24-A06
33	1011	37.9	898	5	BQ649309	BQ649309 AGENCOURT
34	996	37.3	897	5	BU839694	BU839694 AGENCOURT
35	989	37.0	869	8	CX566377	CX566377 UT-M-IBO-
36	988.5	37.0	929	8	CX917126	CX917126 JGI CAN5
37	986	36.9	1001	5	BQ716004	BQ716004 AGENCOURT
38	978	36.6	637	7	CN790575	CN790575 4125180 B
39	971.5	36.4	922	7	CR982195	CR982195 CR982195
40	969.5	36.3	1122	3	BI410402	BI410402 602964915
41	968.5	36.3	756	2	BF134441	BF134441 601784837
42	966	36.2	730	6	CB326067	CB326067 UT-R-D20-
43	964.5	36.1	950	5	BQ648254	BQ648254 AGENCOURT
44	960	36.0	586	6	CD354720	CD354720 UT-M-GMO-
45	956	35.8	819	8	CX907573	CX907573 JGI_CAN4

ALIGNMENTS

RESULT 1	BC068117	3163 bp	mRNA	linear	HTC 08-FEB-2005
LOCUS	BC068117				
DEFINITION	Mus musculus nuclear receptor binding protein 2, mRNA (cDNA clone IMAGE:6419290).				
ACCESSION	BC068117				
VERSION	BC068117.1	GI:45829694			
KEYWORDS	HTC.				
SOURCE	Mus musculus				
ORGANISM	Mus musculus (house mouse)				
REFERENCE	1 (bases 1 to 3163)				
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buettow,K.H., Scheef,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heide,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Donald,M.F., Casavant,T.L., Schenck,T.E., Brownstein,M.J., Utsdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettlemen,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G., Blakesley,R.W., Touchman,D.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalske,U., Smalleg, D.E.,				

Oy		405	oPrcProglucIuvAlaGlnLyAlaAlvThrProthrProglUpProPhaSpSerGluTh	425	
Db		1353	ACCCCAAGGAAGAACCCCAAAGGCCAAATACTCCAAGCGACAACCCTTTGACTGGAGAAC	1412	
Oy		425	TArGLVaLIleGIInetGInCYAsnSLenGluNarSergLunApLySaLaATyrTrpHl	445	
Db		1413	CAGGAAGGTGGTCCAGATGAGTGTCGAACTCGAAMAAGACGAGGACAAAGGCTCGGTGGCA	1472	
Oy		445	bLeuThrLeuleuLeuValIleuGluNapRgYLeuHisArgGlnLeuThryZaaPLeuLe	465	
Db		1473	CCTTAATCTGCTCTTGGCTTGGTAGAACCGGCTCATCGAGCAGCTGCATCTATGATCTGCT	1532	
Oy		465	uPrOThrAspSerXlaGlnNapRLeuAlaSerGlnLeuValHisTryGlyPheLeuHisGl	485	
Db		1593	GGATGACAGCACGACCAAAAGCTAGACGCTTCTGGAAGCACATTTCCTCAAGATACGAGGAGAC	1652	
Oy		505	rGLNALA 507		
Db		1653	GCAAGCG 1659		
RESULT 2					
BC037396					
LOCUS	BC037396	3729 bp	mRNA	linear HTC 05-APR-2005	
DEFINITION	Homo sapiens nuclear receptor binding protein 2, mRNA (cDNA clone IMAGE:5180619), with apparent retained intron.				
ACCESSION	BC037396				
VERSION	BC037396.1	GI:23337035			
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.				
AUTHORS	1. (bases 1 to 3729) Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datschenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stepien, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toehiyuki, S., Carninci, P., Prange, C., Rauba, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mulvihy, S.J., Bosak, S.A., McKean, P.J., McKeaman, K.T., Malek, J.A., Gunaratne, P.H., Richards, S., Weller, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huyl, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Buckson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Beutefield, Y.S., Krzywinski, M.I., Skalka, U., Smalus, D.B., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A. Mammalian Gene Collection Program Team Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932 2. (bases 1 to 3729) . . NIH MGC Project Direct Submission Submitted (04-SEP-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA NIH-MGC Project URL: http://mgc.ncl.nih.gov Contact: MGC help desk Email: gcegbbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc.				
REMARK					
COMMENT					

	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
	DNA Sequencing by: Baylor College of Medicine Human Genome
	Sequencing Center
	Center code: BCM-HGSC
	Web site: http://www.hgsc.bcm.tmc.edu/cdna/
	Contact: amg@bcm.tmc.edu
	Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisgeed, H.,
	Kowis, C.R., Speed, A.J., Martin, R.G., Muzny, D.M., Nannavati,
	A.N., Gibbs, R.A.
FEATURES	
source	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Series: IRAP Plate: 50 Row: b Column: 20 This clone has the following problem: retained intron. Location/Qualifiers 1..3729 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone_image="5180619" /tissue_type="Brain, Lung, Testis, adult, pooled whole" /clone_1ib="Nih MGC_115" /lab_host="DH10B" /note="Vector: pCMV-Sport6"
ORIGIN	
Alignment Scores:	
Pred. No.:	3,74e-205
Score:	2037.50
Percent Similarity:	75.59%
Best Local Similarity:	75.59%
Query Match:	76.31%
DB:	4 Gaps: 4
US-10-618-941-99 (1-507) x BC037326 (1-3729)	
Oy	86 lysillglinthryvalpheluglnleuValaaphisprohenllevallyaleu 105
Db	3 AAATCCAGACCGGTTCGAGCAGCGTGGTGCTGGTGGAACCAACCGAATGTGAAGTTG 62
Oy	106 HlaLySTyTripleuapThSergerlujaCybalary----- 118
Db	63 CACAGTAAGTGGCTGSAITACTTCTGAAGCTGCCGAGGGTGAGACGCGGACAGCCCCG 122
Oy	118 ----- 118
Db	123 GCGAGGACAAGAACACAGCTAGCTGGGGCGGCGCATGGGGAGAACCAGCCGAGCTGTGAA 162
Oy	119 -----VallelhelhellThrgluTryValserSerlySerle 131
Db	183 ATCCTGAGCTCGCACCGTCAGAGTCATCTTCATCAACAAGTAGCTGCATCAGGCACTT 242
Oy	131 uLySGInPhelenuLybLySThTrlyValyaAmhiEllyAlamelcAnlAary----- 148
Db	243 CAAGCAATTCTCTAAAAMAGACCAAGAAAGAACCAAGGCACTGAAACCCCGGTATGGGG 302
Oy	148 ----- 148
Db	303 AGCGGGCTGGGGCAAGCCACGGGGAACAGACGGGGTGGGGGACGCTCGGGGACTGGGATG 362
Oy	148 ----- 148
Db	363 GTGAGGGGGTGCCTGGCGGCTCTCGGACAGGGGCTGGGAGATGCGGGGCGGCGTCCG 422
Oy	149 -----A1 149
Db	423 AGGCCAGCGCCTCTCTGCGGCCCAACCCGACCGAGCGAGTCGTGCGCTCCGCGCGCAGG 482
Oy	149 aTFPLyBATrTPCYSThrglnlleuSerAlaueSerPheleuhIsalaCySerPr 165
Db	483 CTGGAAGCGCTGGTGACAGCAATCTCTGTCCGCTCAGCTTCTTCGACGCGCTCAAGCCC 542
Oy	169 oFolllellEhlsglyAmleuthSerArspThrllebellleghnlbaenglyleu1 169

Db 543 CCAATCATCTCAGGGAAGCTGACGAGCAACCATTTTACAGACACGGCTCAT 602
 Qy 189 elyyllelysevaltrphlsarglllepheserapnallaueuargprothrallaie 209
 Db 603 CAAGATCGGCTCGGTGGCAGCAACCTTCTCCAAAT-----GCACT 644
 Qy 209 uProaBAPleuAisergProllaargalaglwaarglulueuargnleuhspr 229
 Db 645 TCCAGATGATCTCGAAGCCCCATCCGCTGAGCGAGAGAACTTCGAACTTCGACTT 704
 Qy 229 epheroProgluTyrglygluValalaAaerglyThralaValAspillepheserPhegl 249
 Db 705 CTTCCCCCAGAGATGAGAGGCGCGCATGGGACCGCTGTGACATCTTCTCTTGG 764
 Qy 249 yMeCyBAlaLeuGlumeAlaValleuGlullleGlThraNglyAspThraValTh 269
 Db 765 GATGTGTGCTCGAGAGTGGCTGTACTGAAATCCAGACCAATGGGGAACCCGGGTCA 824
 Qy 269 rgluglualalaalargalaaarghisserleuSerAspProaBmeThrgluPhe1 289
 Db 825 AGAGAGGCCATGTGCTGCGCCAGGCACTGCTGAGTGACCCCAACATGCGGAGTTCA 884
 Qy 289 eleuCyBAlaLeuAlaargAspProalaAargProserAlaHisserleuAupheh1 309
 Db 885 CTTTGTGCTGCTGCGCGGACCTTGCCTCCGCGCTCTGCCACAGCTCTCTTCCA 944
 Qy 309 sarvalleuPhegluValHisserleuValleuAlaHisCyBphe1leGlulh1 329
 Db 945 CCCCCTGCTCTTGCAGGTGCATCTGTAAGCTCTTGAGCCCACTGCTTCACTCAGCA 1004
 Qy 329 aglntYrleuMetProgluAaValaValglulYrThraAlaMetAupheh1h1 349
 Db 1005 CCAGTACTTATCTGAGATGTGGAGGAGAGAACCAAGCCATGACCTGACCGC 1064
 Qy 349 aValleuAlaGlueuProargProargProargProleuGlntPargTysergluVal 369
 Db 1065 GGTCTTGCGGAGCTTCCCGGCCCCGAGGCCCTGCTGAGTGGCGGTACTCGAAGT 1124
 Qy 369 lSerPheMetGlueuAuplyAspPheleuGluaPValaArgAaNgly1leTyRProleu 389
 Db 1125 CTCCTTCATGAGCTGAGCAAAATTCCTGAGAGATGTCAAGAAATGGAATCAACCACTGAT 1184
 Qy 389 lAaPheAlaAla1ThraArgProleuGlYleuProargValleuAlaProProProglu 409
 Db 1185 GAACCTTGCAGCACTGACCCCTGGGCTGCCCTGTGCTGCCCCCACCCCGAGAGA 1244
 Qy 409 uValGlntYrAla1YrThraProThraProThraProThraProThraProThra 429
 Db 1245 GGTCCAAAGGCCAAGACCCCGAGCGCAAGGCCCTTTGACTTGAAGACCAAGAGTCA 1304
 Qy 429 eglMetGlntCyAaBleuGluaArgsergluAuplyAaArgTPhlsleuThrlaue 449
 Db 1305 CCAGATGCAATGCAACTGAGAGAGAGCAAGCAAGCGCGCTGGCATCTCATCTGCT 1364
 Qy 449 ulleuValleuGluaAspArgleuH1sarglntleuThraYrAspPheleuProThraAsp 469
 Db 1365 TCTGTGTCTGGAAGACCGCTGCAACCGGACGTCGACTTACGACTCTCCCAACGAG 1424
 Qy 469 rAlaGlntAspLeuAlaSergluLeuValH1sYrGlyPheleuH1sGlu----- 485
 Db 1425 CGCCACGAGACTCGCTCGAGCTGTGACTTGTGCTTCTCTCAGCA-AGTCCGCTGG 1483
 Qy 485 ----- 485
 Db 1484 CGGCGGCGCGGCTGCGGAGGCGGCGAGCGCGGCCCTCGTCCCCCA 1543
 Qy 486 -----AspAaPargMetlyBleuAlaAlaPheleuGlntSerThra 499
 Db 1544 TGCCTCCCTCTTCCGAGAGACACCGAGTGAAGCTGCGCTCTCTGAGAGACACTT 1603
 Qy 499 eleuYrYrArglyThraGlnAla 507
 |||||||

Db 1604 CCTCAAGTACCGGAGGCCAGGCC 1628
 RESULT 3
 AK046142
 LOCUS
 DEFINITION
 AK046142 2163 bp mRNA linear HTC 03-APR-2004
 Mus musculus adult male corpora quadrigenita cDNA, RIKEN
 full-length enriched library, clone: B2J0344117 Product: SIMILAR TO
 NUCLEAR RECEPTOR BINDING PROTEIN (HLR-INTERACTING PROTEIN KINASE)
 homolog [Mus musculus], full insert sequence.
 ACCESSION
 AK046142
 VERSION
 AK046142.1 GI:26337852
 KEYWORDS
 HTC, CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.
 1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>.
URL: <http://fantom.gsc.riken.jp/>.
Location/Qualifiers

FEATURES

source

1. .2163

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM:DB:B230344L17"

/db_xref="taxon:10090"

/clone="B230344L17"

/sex="male"

/tissue_type="corpora quadrigemina"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

/note="unnamed protein product; SIMILAR TO NUCLEAR RECEPTOR BINDING PROTEIN (HLST-INTERACTING PROTEIN KINASE)

homolog [Mus musculus] (SPTK1095045, evidence: FASTY,

97.8%ID, 99.8%length, match=1626)

putative"

/codon_start=1

/protein_id="BAC32612.1"

/db_xref="GI:26337853"

/translation="MSRGESQTVSSGSDPKVSSSLAPGLTSPSPVSTTSAAPE

ERRESEDESEILESPGCRWQKRREYNVPGIDSAVYAMDREBEVEVNVMEVPS

ERNKYLOEKRYAVFDNLQIHLNINVKHYKMAVDYKKAIVIFTEKSSGSKO

FLKTKRHKTKMKEKMKWCTQIQLSLSYLSHCDPILHGNLTCDPIFHNGILKI

GSVPHIRIANVADPTNNHYKTCREBKLNHPAPREGEVNTVTTAVDIYSFMCALB

MAVLEIIOGNSESYVPEQATISSAIQLDEDSLOREFTOKCQSPAPKPAKRELFPA

LFEVSLKLLAAACIVGHOMIPENALEITKMDYSAVIAEIPAGRGEPVQTLYSO

SPALEDKFLDEAVNGIYPLTAGLPKPPQOEVSYPVPSVTPPEPEAEVETR

KVLMOCNISYEVEGVNHLTLTKLEDKRLHLSCLDMNESIPDLAELGOLGIFIS

EADOSRLSYLEETLNKFNRSRSTLNTATVTSVVELT"

2144. .2149

/note="putative"

2163

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

Db

468 CTGAACATCTTAACATTTGTAAGTTTCAACAAATATTGGCTGATGTTAAAGAACAC 527

Qy

117 ALARGValliephelieThrgluYrvalSerSerGlySerLeuYsglnPheLeuYs 136

Db

528 GTAGGGTGATTTTCTATCAGAAATATCTCTGAGGAGCTTAAAGCTTTCTGAAG 587

Qy

137 LyeThrlYelYsAnhiIslYsAlaMetAsnAlaArgAlaTrpYsArgTrpCySerGln 156

Db

588 AAGACCAAAAGAACCAACAGATGAAATGAAAGAGCTTGAAGAGCTGTTGACAG 647

Qy

157 IleuSerAlaLeuSerPheLeuHlaaCySerProProIleIleHlgYsPheLeu 176

Db

648 ATCTCTGCTGCTTAAGCTACCTGACCTCCCTGACCTCCCTCCATCCATGGAACCTG 707

Qy

177 ThSerAspThrIlePheIleGlnHlaaGlnLeuIleYsIleGlySerValTrpHis 196

Db

708 ACCTGACACCACTTCATCCAGACCAAGCACTCATCAAGATTGCTCTGCTTTTCAT 767

Qy

197 ArgIlePheSerAnhiAlaLeuArgProProThrAlaLeuProAspLeuArgSerPro 216

Db

768 AGGATTTTGTCTAAT-----GTGGCTCTGACACTATCAACATCAC 809

Qy

217 IleArgAlaGluArgGluGluLeuArgLeuHlaaPhePheProProGluYrGlyGlu 236

Db

810 GTGAAGCTTGGCCGGGAAGAACAGAACCTTCTTTTGCACACAGATGAGGAA 869

Qy

237 ValAlaAsp---GlyThrAlaValAspIlePheSerPheGlyMetCyValAlaGluMet 255

Db

870 GTACAAACGTGACCAACAGAGTGACATCTACTCTTGGACATGTTGACATGAGATG 929

Qy

256 AlaValLeuGluIleGlnThrAnhiGlySerThrArg---ValThrGluAlaIleAla 274

Db

930 GCAGTCTGAGATTCAGGCGCAATGCGAGTCTCTCATATGCGACAGAGCAATCAGC 989

Qy

275 ArgAlaArgHisSerLeuSerAspProAspMetArgGluPheIleLeuCyValAla 294

Db

990 AGTGCATCCAGCTACTAGAAAGCTATTACAGAGGAGTTTATCAAAGTCCGTCAG 1049

Qy

295 ArgAspProAlaArgArgProSerAlaHlaaSerLeuLeuPheHisArgValLeuPheGlu 314

Db

1050 TCTGAGCTGCTGAGACCAACAGCCAGAACTTGTGTTCCACCAAGCATGTTGAA 1109

Qy

315 ValHisSerLeuYsLeuLeuAlaHlaaHlaaCyPhePheIleGlnHlaaGlnYrLeuMetPro 334

Db

1110 GTGCCCTCACTCAAGCTTCTTGTCTGCTCACTGATATGTTGGGACCAACATGATCCCA 1169

Qy

335 GluAsnValAlaGluGluYrThrYsAlaMetAspLeuHlaaAlaValLeuAlaGluLeu 354

Db

1170 GAGAAAGCTCTAGAGGAGATCCAAAGAACATGATCCAGTCTGTAAGCTGAAT 1229

Qy

355 ProArg---ProArgArgProProLeuGlnTrpArgTrpYsGluValSerPheMetGlu 373

Db

1230 CCCGAGGCGCAGAGCAGAACCAAGTTCAGACTTGTACTCTGACGACCAAGCCCTAGAA 1289

Qy

374 LeuAspIlePheLeuGluGluAspValArgAsnGlyIleTrpProLeuMetAsnPheAla 393

Db

1290 TTGACAAATTCCTTGAAGTGTCAAGATGGAATGGAATCACTCTGACAGCCCTTT 1343

Qy

394 ThrArgProLeuGluYrProArgValLeuAlaProProGluGluVal----- 410

Db

1344 -----GGGTACTGCTGGCCCTCAGACGACACAGAGAGGAGGATCATCACT 1391

Qy

411 -----GlnYsAlaYrThrProThrProGluProPheAspSerGluThrArgYs 427

Db

1392 GTTGTGCCCCCTCTGTCAAGACTCCCAAGCTCCGAGCCAGCTGAAGTGAGACAGAAAG 1451

Qy

428 ValIleGlnMetGlnCyAsnLeuGluYrArgSerGluAspYsAlaArgTrpHisLeuThr 447

Db

1452 GTGTGTGTGATGAGGACATCATCTGTTGAGAGAGGAGACTCAACCACTGATTAACA 1511

Qy

448 LeuLeuLeuValLeuGluAspArgLeuHlaaArgGlnLeuThrYrAspLeuLeuProThr 467

Db

1512 CTTCTGTGAGCTGAGGACAAATTGAACTGGCACTGATCGTGAATGCCAAAT 1571

Db	Accession	Source	Organism	Reference	Authors	Consrtm Title	Journal	Comment
Db	1330	GTGGTGCCTGATGCTGACATGCATCCTGTCGAGGAGGAGGAGTCAACACCATCTAACCA	13899					
Oy	448	Leu1eu1euVal1euG1uAspArg1uHisArg1uLeuThr1uYrAsp1euLeuPro1uThr	467					
Db	1390	CTTCTGCTGAAGCTGACGACAAATGTGAACCGGACCTGAGCTGTGACCTGATGCAAT	14499					
Oy	468	AspSer1uAla1uAsp1uAla1uSer1uLeuVal1uHis1uArg1uPhe1uHis1uGlu1uAsp	487					
Db	1450	GAGGACATCCCGAGCTGGCAGCTGAGCTGGGCGAGCTGGGCTTCAATGAGAGGCTGAT	15099					
Oy	488	ArgMet1uSer1uAla1uPhe1uLeu1uSer1uPhe1uLeu1uYr1uYrArg1uGly1uThr1u	506					
Db	1510	CAGAGCCGCTGACCTCTCTGCTGAGGAGACGCTCAACAGTTCACCTTCAACGAG	1566					
RESULT 5								
LOCUS	HSN805653							
DEFINITION	Homo sapiens mRNA, CDNA DKFZp43412411 (from clone DKFZp43412411).							
ACCESSION	AL334530							
VERSION	AL334530.1							
KEYWORDS	HTC.							
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.							
AUTHORS	1 (bases 1 to 2616)							
CONSRTM TITLE	Bahr, A., Lauber, J., Mewes, H.W., Well, B., Amlid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.							
JOURNAL	The German cDNA Consortium							
COMMENT	Direct Submission Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de; Sequenced by Q1agen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp43412411) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp43412411 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/. Location/Qualifiers 1..2616 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="RZPD:DKFZp43412411" /db_xref="taxon:9606" /clone="DKFZp43412411" /tissue_type="testis" /clone_lib="434 (synonym: htees). Vector pSPORT1; host DH10B; sites NotI + SalI" /dev_stage="adult" /note="hypothetical protein (Rattus norvegicus), N-terminus truncated" 1..2616 /gene="DKFZp43412411" 1..885 /gene="DKFZp43412411" /codon_start=1 /product="hypothetical protein" /protein_id="CAD39186.1" /db_xref="GI:21740357" /db_xref="GO:08NCX8" /db_xref="UniProt/TREMBL:Q8NCX8" /translation="RSP1RARBE1RLN1HPPEYGEVADTAVDIESFGKCALEMAYLEIOTNGDTFTEA1ARARS1LSDPNMRSE1LCLARDPARRSAS1LPHRYLFEVHSLKILNAGT1EHO1YMPENVEVEKTKAM1H1VLA1EPPEPRP1OWYSEVSEFVELDRLE1EVNRN1YPLNMFAR1R1P1GLR1VLA1PPE1EOKAKT1PPEPFDS1RKY1OMOCN1ERS1SD1K1RWH1L1LL1VED1LH1R1OL1TYD1L1P1TSD1AD1L1SELVH1YGF1LHEDDMK1K1A1F1LST1FLK1RG1QA"							

[illegible]

AY408904			1601 bp	DNA	linear	GSS 15-DEC-2000
LOCUS						
DEFINITION	Pan troglodytes NRB gene, VIRTUAL TRANSCRIPT, partial sequence,					
ACCESSION	AY408904					
VERSION	AY408904.1					
KEYWORDS	GSS.					
SOURCE	Pan troglodytes (chimpanzee)					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Pan.					
REFERENCE	1 (bases 1 to 1601)					
AUTHORS	Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smirsky,J.D., Adams,M.D. and Cargill,M.					
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios					
JOURNAL	Science 302 (5652), 1960-1963 (2003)					
PUBMED	14671302					
REFERENCE	2 (bases 1 to 1601)					
AUTHORS	Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smirsky,J.D., Adams,M.D. and Cargill,M.					
TITLE	Direct Submission					
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA					
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.					
FEATURES	Location/Qualifiers					
source	1..1601					
	/organism="Pan troglodytes"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:9598"					
	<1..>.1601					
gene	/gene="NRBP"					
	/locus_tag="HCM3379"					
ORIGIN						
Alignment Scores:						
Pred. No.:	4,51e-152					1601
Score:	1534.50					Matches: 305
Percent Similarity:	73.54%					Conservative: 59
Best Local Similarity:	61.62%					Mismatches: 102
Query Match:	57.47%					Indels: 29
DB:	10					Gaps: 6
US-10-618-941-99 (1-507) x AY408904 (1-1601)						
OY	17 GIUARGUAAPGJUNSERGIUAAPGJUNSERASPILEUGLUGJUNSERPROCGIYATG					36
	123 GGAGAAAGAAGAAAGTAAGTAGTGACTGCAGATTITGGAAAGTCGCCCTGTGGCCOC					182
OY	37 TTPGINLYAARGAUGLUNGIVALAENGINGIYAAMECPROGLIUNGJUNSERTHRPE					56
	189 TCGCAGAAAGGCGAAGAAGGGAATCAACGGAATGTACCAGTATTCACAGTGATCAC					242
OY	57 LEUALAMEASPRTHTGLUGLUGIVAlIGIUIVAITTPANGIULUENHSPHEGLIAP					76
	243 CTGGCCATCATTAACAAGAGGAAGGTGTGAGGTGTGGAATGACGTACAGTTCTGTAG					302
OY	77 ATGYALAphealaaahsiagiugluyisilegiInThValPhegiungInLeuValau					96
	303 CGCAGAACTACCAAGCTCGAGCGAGAAAGGTTCTGCTGTGTATTATATCTGATTCAA					362
OY	97 VALASPHIPROAMNILEVALIYLUENHSLYSTYTPTPLeuaprtHsergiualacy					116
	363 TTGGACACTTTAACTTTGTTAAGTTTCCAATAATTATGGCGTCGATTAAGAACAAG					422
OY	117 AIAAGVALIIepheIIethrgtIyTVaISerSeGIySerIEuylsgInpleuLy					136

Db	423	GCACGGGTCATTTTATTCACAGAAATACATGTCATCTGGAGTGTGAAGCAATTTCTGAG	482
Qy	137	LYSThrIuVlyVaVnHlVlyVaIaMeCAsnIaAArgAlaTrpIyVaLgTrpCySthGln	156
Db	483	AAGACCAAAAGAAACCAAGACCATTAATCAAGACATGGAAGCGTTGGTGCACAA	542
Qy	157	ILeuserAlaLeuserPheLeuHlaIaCySerProPoiIeiHlaGLVaSLeu	176
Db	543	ATCCTCTGCGCTAACTGCTACCTGCTGACCCCTGACCCCATTCATGGAGACTG	602
Qy	177	ThiSerArPThrIlePheIleGlnHlaVaSnIyLeuIleuYsIleGlySerValTrpHis	196
Db	603	ACCTGTACACCACTTCATTCACACACAGCATTCACAAAGATTGGCTGTG-----	656
Qy	197	ArgIlePheSerAsnAlaLeuArgProProThrAlaLeuProArPheLeuArgSerPro	216
Db	657	-----GCTCGTCACTATTCACAAATCAT	680
Qy	217	ILeargalAGIuArgGluGluLeuArgAsnLeuHisPhePheProProGluIuYrGlyGlu	236
Db	691	GTGAAGACTTGTGAGAAAGACAGAAATCTTACACTTTTGACACCAAGATATGAGAA	740
Qy	237	ValAlaAsp---GlyThrAlaValAspIlaPheSerPheGlyMetCyVaIaLeuGlyMet	255
Db	741	GTCACTATATGTGACAAACAGCAGTGCATCTACTCTTGGCAGTGTGATGCATGGAATG	800
Qy	256	AlaValleuGluIleGlnThrAsnGlyVaSPThrArg--ValThrGluAlaIleAla	274
Db	801	GCAGTGTGAGATTCACGGGCAAATGGAGAGACTTCATATGTGCCACAGAAAGCATAGC	860
Qy	275	ArgIlaaArgHisSerLeuserSerArPAsnMetArgGluPheIleLeuCyVaLeuAla	294
Db	861	AGTGCATCCAGCTTTAGAAAGCCCATTCAGAGGAGAGTTCATTCAAAAGTGCCTGAG	920
Qy	295	ArgArPProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGlu	314
Db	921	TCTGAGCTGTCTCCGACACCAACAGCCAGAAATCTTGTTCCACCACAGATGTTTGAA	980
Qy	315	ValHisSerLeuLyLeuLeuAlaAlaHisCySerPheIleGlnHisGlnIuYrLeuMetPro	334
Db	981	GTGCGCTGTGCAAACTCTTGCGGCGCCACTGTGTGGGACACCAACATGATCCCA	1040
Qy	335	GluAsnValValGluGluYsThyIyAlaMetAsPheuHisAlaValIleuAlaGluLeu	354
Db	1041	GAGAAACGCTCGAGAGAGATCACCAAAAACATGGATCTACTGCTGTGATCGGCTGAATC	1100
Qy	355	ProArg---ProArgArgProLeuGlnIuTrpArgIuYrSerGluValSerPheMetGlu	373
Db	1101	CTCGACAGACCAAGAAAGAAACCAAGTTCAAGCTTTGATCTTCAGTACCAAGCTCTGAA	1160
Qy	374	LeuAspLyPheLeuGluAspValArgAsnGlyIleTyProLeuMetAsnPheAlaIa	393
Db	1161	TTAGATTAAATTCCTTGAAGATGTCAAGGATGGATCTATCCTTGACAGCTTT-----	1214
Qy	394	ThrArgProLeuGlyLeuProArgValLeuAlaProProGluGluVal-----	410
Db	1215	-----GGGCTGCTCGGCGCCGACGACGACAGCAGAGAGAGAGATACATCAGCT	1262
Qy	411	-----GlnlyVaIaYsThProThrProGluProPheAsnSerGluThArgYs	427
Db	1263	GTGCGGCCCTCTGTGACATCTCAACACTTAACCAAGCTGAGGTGGAATCTGCAAG	1322
Qy	428	ValIleGlnMetGlnCyAsnLeuGluArgSerGluAspLyValArgTrpHisLeuThr	447
Db	1323	GTGGTGTGATGACAGTGCACATTTAGTGGGTGAGAGGAGGAGCAAAACCACTGAGC	1382
Qy	448	LeuLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThrTyTrpArPheLeuProThr	467
Db	1383	CTTCTGCTGAAGTTGGAGCAAACTGAACCGGACCTGACCTGTGACTATGTCCAAT	1442
Qy	468	AspSerAlaGluAspLeuAlaSerGluLeuValHisTyTrgIyPheLeuHisGluAspAr	487
Db	1443	GAGATATATCCCGAGGTGGCGGCTGAGCTGGGTGAGAGCTGGGCTTCATTATAGTGAAGCTGAC	1502

QY 295 ArgAspProAlaArgArgProSerAlaHisSerLeuPheHisArgValLeuPheGlu 314
Db 1015 TCTGAGCTCTCTGCGACGACCAACAGCCAGAACTCTGTTCAACCCAGCATTTGTTGA 1074
QY 315 ValHisSerLeuValSerLeuValAlaHisCysPheIleGlnHisGlnTyrLeuMetPro 334
Db 1075 GTCCCTCGCTCAAACTCTTGCGGCCCACTGCATTTGTGGAGACCAACAATGATCCCA 1134
QY 335 GluAsnValValIleGluTyrThrValAlaMetAspLeuHisAlaValLeuValGluLeu 354
Db 1135 GAAACAAGCTCTTAAGAGGATCACCAAAACATGAATACTAGTCCCTGACCTGGAATTC 1194
QY 355 ProArg---ProArgArgProLeuGlnTyrArgTyrSerGluValSerPheMetGlu 373
Db 1195 CTGCGAGAGACCAAGAAAGAACAGATTCAAGCTTTGATCTCTCACTCAAGTCCAGAA 1254
QY 374 LeuAspLysPheLeuGluAspValArgAsnGlyIleTyrProLeuMetAsnPheAla 393
Db 1255 TTGATTAATATCTCTTGAAGATGTCAGGAATGGGATCATCTCTGACAGCCTTT----- 1308
QY 394 ThrArgProLeuGluLysProArgValLeuAlaProProGluGluVal----- 410
Db 1309 -----GGGCTGCTCGGCCCCAGCACGCCACAGCAGAGAGGTGACATCACCT 1356
QY 411 -----GlnLysAlaLysThrProThrProGluLysProPheAspSerGluThrArgLys 427
Db 1357 GTCTGCCCCCTCTCTCAAGACTCCGACCTGAACTGAACTGAGTGAAGTCCCAAG 1416
QY 428 ValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLysAlaArgThrPheIleThr 447
Db 1417 GTGGTCTGATGACAGTGCAACATTTAGTGGTGAAGAGGAGATCAACACCCCTGAC 1476
QY 448 LeuLeuLeuValLeuGluAspArgLysHisArgGlnLeuThrTyrAspLeuLeuProThr 467
Db 1477 CTCTCTCTGAAGTTGAGGACCAAACTGAACCGGACCTGACCTGATGAGCCCAAT 1536
QY 468 AspSerAlaGlnAspLeuValSerGluLeuValHisTyrGlyPheLeuHisGluAspArg 487
Db 1537 GAAATATATCCCGAGTTGGGGCTGAGCTGTGACGTGGCTTCAATTAGTGAAGCTGAC 1596
QY 488 ArgMetLysLeuAlaAlaPheLeuGluSerThrPheLeuLysTyr 502
Db 1597 CAGAGCCGGTTGACTTCTCTGTAAGAGACCTTGAACAAGTTC 1641
RESULT 10
CR858669 2181 bp mRNA linear HTC 12-NOV-2004
LOCUS Pongo pygmaeus mRNA, cDNA DKFZp459G2032 (from clone DKFZp459G2032).
DEFINITION CR858669
ACCESSION CR858669
VERSION CR858669.1 GI:55728271
KEYWORDS HTC
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Pongo.
REFERENCE 1 (bases 1 to 2181)
AUTHORS Koehler, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
CONSTRM Fobbo, G., Han, M. and Wiemann, S.
TITLE The German cDNA Consortium
JOURNAL Direct Submission
COMMENT Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the
Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA
Sequencing Consortium of the German Genome Project. This clone
(DKFZp459G2032) is available at the RZPD Deutsches
Resequenzierungszentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp459G2032
Further information about the clone and the sequencing project is

FEATURES available at <http://mips.gsf.de/projects/cdna/>.
source location/Qualifiers
1..2181
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp459G2032"
/tissue_type="cortex"
/clone_lib="459 (synonym: poor1). Vector pSport1_Sfi; host
DH10B; sites SfiI + SfiIB"
/dev_stage="adult"
/note="nuclear receptor binding protein (Homo sapiens)"
1..2181
/gene="DKFZp459G2032"
116..1723
/gene="DKFZp459G2032"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH90881.1"
/db_xref="GI:55728272"
/translation="MSGESQTVLSGSDPKYSSSSAPGLTSPVPTVSTSAAPS
EESESDSEILIESPCRMQRREVNQRPVGDSDAYLADTREGVENVNVS
ERKNYKIQEKVAVFDNLIOLEHINIYKFKYVADIKENKARVIFITETMSGLIKI
FLKTKRKHKTWNEKAMKRMCTQILSALSYHSQDPIHGMLTCDTPIHNGLIKI
GSVADDTNNHYKTREROKNHPAPRYGEVNTYTTVDIYSPGCALEMVLTEOG
NGESSYVPEALISATIQLEBDLQREPFQKCIQSERPARPTRELLFHPALREVSLEK
LTLAHCIVGHQMTPEMLBEITKMDTSVAVALIIPAGGRPPVOTLVISQSPALELDK
FLEDRVNGIYPLTAPRPPQPOQPEETSPVPSVKPTPEPRAVEVTRKVLQCN
ISVEGVGHHTLLTKLEDKELKRLHSQCLMPEENIPELAEVLQGFISEADQSLT
SLBERTLNKFNPARSSTLSNAVTVSS"
ORIGIN
Alignment Scores:
Pred. No.: 7.27e-152 Length: 2181
Score: 1534.50 Matches: 305
Percent Similarity: 73.54% Conservative: 59
Best Local Similarity: 61.62% Mismatches: 102
Query Match: 57.47% Indels: 29
DB: 4 Gaps: 6
US-10-618-941-99 (1-507) x CR858669 (1-2181)
QY 17 GluArgGluAspGluSerGluAspGluSerAspIleLeuGluGluSerProCysGlyArg 36
Db 245 GAGGAAGAAGAAAGAAAGTGAAGATGATGCTGTGAGATTGGAAAGAGTCGCCATGTGGCCG 304
QY 37 TPGAlnLysArgArgGluGlnValAlaGlnGlyAsnMetProGlyLeuGlnSerThrPhe 56
Db 305 TGGCAGAAGAGCGGAGAAGAGTGAATCAACGGAATGTACAGATTTACAGCTGATAC 364
QY 57 LeuAlaMetAspThrGluGluGluValAlaValAlaTTPAsnGluLeuHisPheGlyAsp 76
Db 365 CTGGCCATGATATACAGAGGAAGGTAGAGTGTGTGGAAAGAGGTACAGTTCTCTGAG 424
QY 77 ArgLysAlaPheAlaAlaHisGlnLysIleGlnThrValPheGluGlnLeuValLeu 96
Db 425 CGCAAGACTAACACCTGCAGAGGAAAAAGTTGCGCTGTGTGTGATATATCGATTACG 484
QY 97 ValAspHisProAsnIleValLysLeuHisLysTyrThrLeuAspThrSerGluAlaCys 116
Db 485 TTGGAAGATCTTAACATTTGTAAGTTTCACAAAATCTGGGCTGACATTAAAGAAACAAG 544
QY 117 AlaArgValIlePheIleThrGluTyrValAlaSerSerGlySerLeuLysGlnPheLeuLys 136
Db 545 GCCAGGTCATTTTATATCAACAATATACATCATCTGGAGTCTGAAGCAATTTCTGAAG 604
QY 137 LysThrLysLysAsnHisLysValAlaMetAsnAlaAlaGlaTTPLysArgTTPCysThrGln 156
Db 605 AAGACCAAAAMAAACAACAAGCATGATGAAAAAGGCAATGAAGGTTGGTGCACACA 664
QY 157 IleLysSerAlaLeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAsnLeu 176
Db 665 ATCTCTCTGCCCCCTAAGTCACTGACCTCTGTGACCCCCCATCATCATGAGGAACCTG 724

[illegible]

	ACCESSION	CRS91118	of Homo sapiens (human) .
	VERSION	CRS91118.1 GI:50471925.	
	KEYWORDS	HTC; CNSIT_CDNA.	
	SOURCE	Homo sapiens (human)	
	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
	REFERENCE	1 (bases 1 to 2086) Li,W.B., Gruber,C.J., Jesssee,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished	
	AUTHORS	Contact : Feng Liang Email : fliang@lifetech.com URL :	
	JOURNAL	http://fulllength.invitrogen.com/Invitroden Corporation 1600 Paradise Avenue Foster City CA 94024 Genoscope. Direct Submission Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr Web site: www.genoscope.cns.fr The sequence was generated by PCR amplification from total RNA extracted from HeLa cells transfected with NotI oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies , a division of Invitrogen. Location/Qualifiers 1..2086 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /cisname="CSODK009YM06" /rissue type="HeLa cells Cot 25-normalized" /plasmid="pCMVSPORT_6"	
	FEATURES	source	
	ORIGIN		
	Alignment Scores:		
	Pred. No.:	2.39e-150	Length: 2086
	Score:	1520.00	Matches: 303
	Percent Similarity:	73.54%	Conservative: 61
	Best local similarity:	61.21%	Mismatches: 101
	Divergence Match:	56.93%	Gaps: 30
	DB:	4	Indels: 7
	US-10-618-941-99 (1-507) x CRS91118 (1-2086)		
OY	17	GLUARGLUAPGILUSERGIUAAPGILUSERASPILEUGLIGUSERPROCYSGIAYRG	36
	:	:::::	:---
Bd	214	GAGAGAAMAAAGATGAAGTGCTCGAATTGGAMAAGTCGCCCTGTGGCCGC	273
OY	37	TTRGTLNLYARGARGTGUGLVLAANGNLGYASMEEPRCPLJENGINSERTINPHE	56
Bd	274	TGGCAAGAACGGCGAGAGAGGTGAATCAAAGAAATGACCAGTAGTAACAAGTCATAC	333
OY	57	LNUALAWECAPRTHNGIUGLIVYLGUVAVAI TRPAANGILUNEUHIAPHEGLIASP	76
Bd	334	CTGGCATGTGATA CAGAGGAAGGTGRNAAGTTGTGTGGAATGAGGTACACTTCTCGAA	393
OY	77	ARYLVREIALPHELALAHISGLIGLVLVSILEGINTRYALPHEGELGUNLEVAILLEU	96
	:	:::	:---
Bd	394	CGCAAGAACACACAAAGCTGCAGAGGAAAAGTTCCGTGCTGTGTGATATCATGATCAA	453
OY	97	VALLSPHISPRAANTLEVALVLSLEUHII SYRYTYRTPLEUBAPHHSERGIUALACYG	116
	:	:::::	:---
Bd	454	TTGAGAGATTCTTAACATGTGTAAGTTTCCAATAATATGGGCTGACATTAAGAGAACAG	513
OY	117	ALAARGVAIIIIEPHELLERTHGILURYVALSERSECIYSERYNEUYSGINPEULEUYS	136
Bd	514	GCCAGGGTATTTTTATTCACAGAAATACATGTATCTCGGAGICTGAAGCAATTTCTGMA	573
OY	137	LYETHLLYLVEKAHHISLYVALAWECUANAL AARGALETRPLYAARGTRPCYSTHRGIN	156
	:	:::::	:---

Db	574	AAAGCCAAAAGAACCAACAAGCATGAATGAATAAAGGATGGAGCCGTTGGTGCACACAA	633
Qy	157	IlleuSerAlaLeuSerPheLeuH1ValAcySerProProIleH1GlyValLeu	176
Db	634	ATCCCTCTGCGCCTTAAGCATCTGCAGCTCTGTGAACCCGCCATCATCCATGGAACTGG	693
Qy	177	ThrSerArgThrIlePheIleGlnH1ValGlnGlyLeuIleValIleGlySerValTrpH1s	196
Db	694	ACCTGTGCACACATCTTTCATCCAGACAAACGAGATCATCAAGATTGGCTGTGG-----	747
Qy	197	ArgIlePheSerAspAlaLeuArgProProThrAlaLeuProArgPheLeuArgSerPro	216
Db	748	-----GCTCCGACCACTTAATCAATCAT	771
Qy	217	IleArgAlaGluArgGluLeuArgValLeuH1SerPheProProGluTrpGlyGlu	236
Db	772	CTGGAAGCATCTGTCCAGAGAGACAGAAAGATCTACATCTTTGGACCCAGAGATGTGAAGA	831
Qy	237	ValAlaArg--GlyThrAlaValArgIlePheSerPheGlyMetCysValAlaLeuGlyMet	255
Db	832	GTCACTAATGTACAACAGACATGGACATCTACTCTTGGCATGTGGCATGGAGATGG	891
Qy	256	AlaValLeuGluIleGlnThrArgGlnArgTrpArg--ValThrGluGluAlaIleAla	274
Db	892	CGAGTGTCTGAATTCACGGCAATGAGAGTCTCATATGTGTCACAGAAACCCATCAGC	951
Qy	275	ArgAlaArgH1SerLeuSerArgProArgMetCysGluPheIleLeuCysCysValLeuAla	294
Db	952	ACGTCCATCCACATCTTGAAGAACCA-----TTACAGAGTTTATTCAAAAGTCCGACG	1008
Qy	295	ArgArgProAlaArgArgProSerAlaH1SerLeuLeuPheH1ArgValAlaLeuGlu	314
Db	1009	TCTGAGCCTGCTCCGACACCAACAGCAGAGAACTTCTGTTCCACCCAGCATTTGTTGAA	1068
Qy	315	ValH1SerLeuLeuLeuAlaH1ValH1SerPheIleGlnH1GlyTrpLeuMetPro	334
Db	1069	GTGCGCTGCTCAAACTCTTCCGGCCCATCTGATGTGGACACCAACATGTATGCCA	1128
Qy	335	GluArgValValGlnGluTrpThrArgValMetCysPheH1AlaValAlaLeuArgLeu	354
Db	1129	GAGAAACGCTCTAGAGAGATGACCAAAAACATGGATCTAGAGCCGTACTGGCTGAATTC	1188
Qy	355	ProArg--ProArgArgProProLeuGlnTrpArgTrpSerGluValSerPheMetGlu	373
Db	1189	CTCGACAGACACAGAAAGAGAACAGATCCAGATTTGTACTCCAGTCCACACAGCTGGAA	1248
Qy	374	LeuArgArgPheLeuGlnArgValAlaArgArgGlyIleTrpProLeuMetAspPheAlaAla	393
Db	1249	TTAGATTAATTCCTTAGAGATGTCCAGAAATGGGATCTTACCTCTACAGCCTTT-----	1302
Qy	394	ThrArgProLeuGluLeuProArgValIleAlaProProProGluGluVal-----	410
Db	1303	-----GGGTGCTGCTGGCCCCAGACGCCACAGAGAGAGAGATGACATCAGCT	1355
Qy	411	-----GlnValAlaLeuThrProThrProGluProPheArgSerGlnThrArgGly	427
Db	1351	GTGCTGCCCCCTGTGTGAAGATCCGACACCTGAACCAACGCTGAGGTGAGATGGTGAAG	1410
Qy	428	ValIleGlnMetGlnCysValLeuGluTrpSerGlnArgArgValAlaArgTrpH1ValLeuThr	447
Db	1411	GTGCTGTGATCCAGTGCACAACTTGATGTGCTGAGAGGGAGAGTCAACACCACTGTGCA	1470
Qy	448	LeuLeuLeuValLeuGlnArgArgArgValArgGlnLeuThrTrpArgLeuLeuProThr	467
Db	1471	CTTCTGTGCTGAAGTGGAGACAACTGAACCCGACACCTGAGCTGTGACTGATGTGCAAT	1530
Qy	468	AspSerAlaGlnArgPheLeuLaseGluLeuValH1SerGlyPheLeuH1GlyValLeuArg	487
Db	1531	GAGAAATATCCCGAGATGGCGGCTGAGCGTGGTGCAGAGTGGGCTTCAATTAAGTAGGCTGAC	1590
Qy	488	ArgMetLeuValLeuAlaPheLeuGluSerThrPheLeuLeuTrpTrp	502
Db	1591	CAGACCCGGTGTGATCTTCTGTAGAAAGACCTTGAACCAAGTTC	1635

RESULT 12	AKO77440	2181 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AKO77440				
DEFINITION	Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched library, clone:5730409G3 product:SIMILAR TO NUCLEAR RECEPTOR BINDING PROTEIN (HUS7-INTERACTING PROTEIN KINASE) homolog [mus musculus], full insert sequence.				
ACCESSION	AKO77440				
VERSION	AKO77440.1	GI:26097350			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryotic, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Muridomorpha; Muridae; Murinae; Mus.				
AUTHORS	1 Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Suganara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitanishi, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikake, S., Inoue, K., Togawa, Y., Izawa, M., Onate, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
PUBMED	11076861				
REFERENCE	4				
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409, 685-690 (2001)				
PUBMED	5				
REFERENCE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.				
AUTHORS	Analysis of mouse transcriptome based on functional annotation of 60,770 full-length cDNAs				
TITLE	Nature 420, 563-573 (2002)				
JOURNAL	6 (bases 1 to 2181)				
REFERENCE	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, Y., Toya, T., Yasunishi, A., Yumatsugu, M., and Hayashizaki, Y.				
AUTHORS	Muramatsu, M. and Hayashizaki, Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-APR-2002) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Chemical Research, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)				
COMMENT	cDNA library was prepared and sequenced in Mouse Genome				

VERSION	CR998084.1	GI:68292689
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 899)	
AUTHORS	Heil,O., Ebert,U., Hennig,S., Henze,S., Radelof,U., Schneider,D. and Korn,B.	
TITLE	Human T-Lymphocytes library	
JOURNAL	Unpublished (2005)	
COMMENT	Contact: Inge Ariart RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Heubnerweg 6, D-14059 Berlin, Germany Email: www.rzpd.de RZPD: RZPDp9017B1314. RZPDLIB: (Human T-Lymphocytes) RZPD LIB No.9017 http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact: Inge Ariart RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Heubnerweg 6, D-14059 Berlin, Germany Tel: +49 30 32639 100 Fax: +49 30 32639 111 www.rzpd.de This clone is available from RZPD: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=RZPDp9017B1314 contact RZPD (product-support@rzpd.de) for further information. Primer name: qe3.4, Primer sequence: CGGATTAACAATTCACACAG. location/Qualifiers 1..899 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="RZPDp9017B1314" /tissue_type="T-Lymphocytes" /dev_stage="adult" /lab_host="DH10B" /clone_lib="RZPD no.9017" /notes="vector: pQE80LSN_cloned; site_1: SalI; site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned.pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACGTCTGATGATCGCAGCGCCGCTTTTCTTTTCTTTT 3'] Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"	
ORIGIN		
Alignment Scores:		
Pred. No.:	9,4e-147	length: 899
Score:	1481,00	Matches: 290
Percent Similarity:	96,31%	Conservative: 1
Best Local Similarity:	97,97%	Mismatches: 4
Query Match:	55,47%	Indels: 2
DB:	7	Gaps: 0
US-10-618-941-99 (1-507) x CR998084 (1-899)		
QY	212	AppleusarSerpIleArGlaGluArgGluGluLeuArgAsnLeuHisPhePhePro 231
DB	1	GATCTCCGAAGCCCATCCGCGCTAGGCGAGAGAACTTCGGAACCTGCATCTTCTCC 60
QY	232	ProGluTYGlyGluValAlaAspGlyThrAlaValAspIlePheSerPheGlyMetCys 251
DB	61	CCAGAGTATGGAGAGGTGGCCGATGGACCCGCTGTGGACATCTTCTCTTGGGATGTGT 120
QY	252	AlaLeuGluMetAlaValLeuGluIleGlnThrAsnGlyAspThrArgValThrGluGlu 271
DB	121	GGCGTGGAAATGAGCTGATCGGAATCCAGACCAATGGGGAACCCGGGGTCAAGAGAG 180
QY	272	AlaIleAlaArgAlaArgHisSerLeuSerLeuAspProAsnMetArgGluPheIleLeuCys 291

Db	LOCUS	DEFINITION	ACCESSION	VERSION	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	BX410189	RESULT 14	BX410189	BX410189.1	GI:30768891	Homo sapiens (human)					
Db	BX410189	LOCUS	BX410189	BX410189.1	GI:30768891	Homo sapiens (human)					
Db	BX410189	DEFINITION	BX410189	BX410189.1	GI:30768891	Homo sapiens (human)					
Db	BX410189	ACCESSION	BX410189	BX410189.1	GI:30768891	Homo sapiens (human)					
Db	BX410189	VERSION	BX410189	BX410189.1	GI:30768891	Homo sapiens (human)					
Db	BX410189	SOURCE	BX410189	BX410189.1	GI:30768891	Homo sapiens (human)					
Db	BX410189	ORGANISM	BX410189	BX410189.1	GI:30768891	Homo sapiens (human)					
Db	BX410189	REFERENCE	BX410189	BX410189.1	GI:30768891	Homo sapiens (human)					
Db	BX410189	AUTHORS	BX410189	BX410189.1	GI:30768891	Homo sapiens (human)					
Db	BX410189	TITLE	BX410189	BX410189.1	GI:30768891	Homo sapiens (human)					
Db	BX410189	JOURNAL	BX410189	BX410189.1	GI:30768891	Homo sapiens (human)					
Db	BX410189	COMMENT	BX410189	BX410189.1	GI:30768891	Homo sapiens (human)					

```

SOURCE
1. .927
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN002Y109"
/cisue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_idb="Homo sapiens ADULT BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer; five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ALIGNMENT SCORES:
Pred. No.: 2.06e-138 Length: 927
Score: 1403.00 Matches: 275
Percent Similarity: 94.24% Conservative: 3
Best Local Similarity: 93.22% Mismatches: 15
Query Match: 52.55% Indels: 2
DB: Gaps: 0

US-10-618-941-99 (1-507) x BX410189 (1-927)

QY 214 ArgSerProIleArgAlaGluArgGluLeuArgAsnLeuHisPheProProGlu 233
Db 3 CGAAGCCCATCCCGCTGACCGAGAGAACTTCGAACTGCACTTCTCCCGCCAG 62
QY 234 TyTgYgUuValaAaepgYThraValaAspIlePheSerPheGlyMetCysAlaLeu 253
Db 63 TATGGAAGGTGGCGCATGGGACCGCTGTGGACATCTTCTTGGAGTGTGGCGCT 122
QY 254 GluMetLalaValLeuGluIleGlnThraAsnGlyAspThrArgValThrGluAlaIle 273
Db 123 GAAGTGGCTGTACTCGAAATCCAAACCAATGGGACACCGGGTCCACAGAGGCGCAT 182
QY 274 AlaArgAlaArgHisSerLeuSerAspProAsnMetArgGluPheIleLeuCysValLeu 293
Db 183 GCTCGCCCGCAGGCACTGCTGAGGACCCCAACATGGGAGTTCACTTGTGCTGCT 242
QY 294 AlaArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValaLeuPhe 313
Db 243 GCCCGGGACCTCGCGCGCGCTCTGCCACAGCTCTCTTCCACCGCGTCTTTC 302
QY 314 GluValHisSerLeuValLeuAlaAlaHisCysPheIleGlnHisGlnTyrLeuMet 333
Db 303 GAGGTGCACTCGCTGAAGCTCTGCGCAGCCCACTGCTTCACTCAGCACAGTACCTCATG 362
QY 334 ProGluAsnValaValGluGluValThrValaMetAspLeuHisAlaValaLeuAlaGlu 353
Db 363 CCGAGAGATGTGTGGAGGAGAAAGCAACAGCCATGTGACCGGGTCTTGGCGGAG 422
QY 354 LeuProArgProArgArgProProLeuGlnTTPArgTyrSerGluValSerPheMetGlu 373
Db 423 CTTCGCCGCGCGCGCGCGCGCGCTGCACTGCGAGTCTCTTCACTGAG 482
QY 374 LeuAspIlePheLeuGluAspValaArgAsnGlyIleTyrProLeuMetAsnPheAlaIle 393
Db 483 CTGAGCAAAATTCCTGGAGATGTCAAGAAATGAAATCACTGAACTTGGAGGCC 542
QY 394 ThrArgProLeuGluValProArgValaLeuAlaProProProGluValaGlnVala 413
Db 543 ACTCGAACCTCGAGGCTGCGCGCTGTCTGCGCCCAACCGAGGAGGTCGCAAAAGGCC 602
QY 414 LysThrProThrProGluProPheAspSerGlyThrArgValaIleGlnMetGlnCys 433
Db 603 AAGACCCCGACGCCAGAGCCCTTGAATCTGAGACCGAAGGTCACTCAAGATGCACTGC 662
QY 434 AsnLeuGluArgSerGluAspValaValaArgTyrPheIleThrLeuLeuValaLeuGlu 453
Db 663 AACCTGGAGAGAGCGAGCAAGCGCGCTGCGCATCTCACTGCTTCTGAGCTTGA 722

```

```

QY 454 -AspArgLeuHisArgGlnLeuThrTyrAspLeuLeuProThrAspSerAlaGlnAspLe 473
Db 723 AGACGGCTCACCCGANNAGTGAACCTTACGAACTGTCCACCGGACAGGCCCTCAAGACT 782
QY 473 uAlaSerGluLeuValHisTyrGlyPheLeuHisGluAspAspArgMetLysLeuAlaI 493
Db 783 CGCCTTGAGAGCTCGGGGACATATGAGCTTCTCCACAGAAACACCGGTATGAAGCTGGCCG 842
QY 493 aPheLeuGlu-SerThrPheLeuLysTyrArgGlyThrGln 506
Db 843 CTTTCTGAAAGGCCCTTCTTAAGTCCCGGAGGACCCAG 883

RESULT 15
B1645887/c 825 bp mRNA linear EST 12-SEP-2001
LOCUS DEFINITION 603275555F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5315916 5',
mRNA sequence.
ACCESSION B1645887
VERSION B1645887.1 GI:15560123
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 825)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strauberg, Ph.D.
Email: cgaap-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LLNL1800 Row: n column: 13
High quality sequence stop: 808.
Location/Qualifiers
1. 825
source
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129, C57BL/6J, FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5315916"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_note="DH10B"
/clone_idb="NCI_CGAP_Mam3"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
```

ORIGIN

```

ALIGNMENT SCORES:
Pred. No.: 6.21e-120 Length: 825
Score: 1229.00 Matches: 239
Percent Similarity: 92.86% Conservative: 8
Best Local Similarity: 89.85% Mismatches: 11
Query Match: 46.03% Indels: 8
DB: Gaps: 1

US-10-618-941-99 (1-507) x B1645887 (1-825)

QY 140 LysAsnHisLysLeuAlaMetAsnAlaArgAlaTTPLYAspTyrCysThrGlnIleLeuSer 159
Db 776 CGACGAGTGTGTCGCTCCGCTGCGCAGGCCCTGGAACCGCTGTGATCGAGATCTGTCT 717

```

```
QY 160 AlaLeuSer-PheLeuHisAla-CysSerProProIleIleHisGluAsnLeuThrSerA 179
Db 716 GCACTCAGTTTTTTGGACGCTCCGACATCATCGGAGACCTGACGAGCG 657
QY 179 sPTxrlIephelIegInHisAangLYLeuIleLysIleGlySerValTyrPHisArgIleP 199
Db 656 ACACCATCTTCACTTACGACCAATGGCTCATCAAGATCGGCTCGTGTGTACCGCATCP 597
QY 199 heseRenalaleuArgProProThrAlaleuProAspAspLeuArgSerProIleArgA 219
Db 596 TCTCCCAAT-----GCACCTTCTGTATGATCTCCGAGGCCATTATTCGAG 555
QY 219 laGluArgGluGluLeuArgAsnLeuHisPhePheProGluTyrGluValAlaA 239
Db 554 CTGACGGGAGAACCTCGAACCTGCACTTTTCCACCAAGATATGGCAAGTCAATG 495
QY 239 sPGLYThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaValLeug 259
Db 494 ATGGGACTGTGTGGACATCTTCTCCTTGGGATGTGTGCACTGGAGATGGCTGTACTCG 435
QY 259 luIleGlnThrAangLYAspThrArgValThrGluGluAlaIleAlaArgAlaArgHis 279
Db 434 AGATCCAGGCCAACGGGGATACCAAGATCAAGAAAGGCCATGCTCGAGCCAGGCACCT 375
QY 279 erLeuSerAspProAspMetArgGluPheIleLeuCysCysLeuAlaArgAspProAlaA 299
Db 374 CACTGAGTGACCCCAATGCGGAAATCATCTCTGCTGCTGGCCCGGACCTGCCCC 315
QY 299 rGARProSerAlaHisSerLeuLeuPheHisArgValLeuPheGluValHisSerLeuL 319
Db 314 GCCGACCTTCAGCCCAACACTCTCTTCCACCGAGTGTCTTTGAGTGACACTCGCTGA 255
QY 319 yLeuLeuAlaAlaHisCysPheIleGlnHisGlnTyrLeuMetProGluAsnValValG 339
Db 254 AGCTGCTGGCAGCTCACTGCTTCATCCAGCACCACTCATGCTGAGAAATGTGTAG 195
QY 339 luGluLysThrLysValMetAspLeuHisAlaValLeuAlaGluLeuProArgProArgA 359
Db 194 AGGAAAGACCAAGGCCATGACCTTCATGCACTTTGGCTGAGATGCCGAGCCCCCATG 135
QY 359 rGProProLeuGlnTTPArgTyrSerGluValSerPheMetGluLeuAspLYPheLeug 379
Db 134 GACCCCAATGCACTGGCGGTACTCAGAGGTCTCTTCTTGGAGCTGGACAAATTCCTAG 75
QY 379 luAspValArgAangLYIleTyrProLeuMetAsnPheAlaAlaThrArgProLeuGlyL 399
Db 74 AGGATGTCAGAGAACGGGATCTATCCACTGATGAACCTTGGCGGTGCTCGGCCCTTGGGGC 15
QY 399 euProArgValLeu 403
Db 14 TTCCCCGTGTGTG 1
```

Search completed: January 14, 2006, 00:46:08
Job time : 4643 secs


```
RESULT 2
US-09-862-027-10
; Sequence 10, Application US/09862027
; Patent No. 6858418
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. 6858418el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(326)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-862-027-10

Query Match          59.2%; Score 1580; DB 2; Length 326;
Best Local Similarity 97.4%; Pred. No. 1,2e-145;
Matches 301; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

QY 1 MAPEPAPRAREREREDESEDESDIIESEPCGMRQREOVNQGNNPGLQSTFLAMD 60
    |||||
DB 12 MAPEPAPRAREREREDESEDESDIIESEPCGMRQREOVNQGNNPGLQSTFLAMD 71
    |||||

QY 61 TEEGVVNNELHFGDRKAPAAHEKIQTFEQLVVDHPNIVKLHKYLDTSEACARYI 120
    |||||
DB 72 TEEGVVNNELHFGDRKAPAAHEKIQTFEQLVVDHPNIVKLHKYLDTSEACARYI 131
    |||||

QY 121 FITEVSSGSLKQFLKTKKNHAKANARAKMCTQIISALSFLLHACSPPIIHGNTSDT 180
    |||||
DB 132 FITEVSSGSLKQFLKTKKNHAKANARAKMCTQIISALSFLLHACSPPIIHGNTSDT 191
    |||||

QY 181 ITIQHNGLIKIGSVWHIRFISNAPPTALPDDLRSPIRARREELRNLFHPPRYGEVADG 240
    |||||
DB 192 ITIQHNGLIKIGSVWHIRFISN-----ALPDLRSPIRARREELRNLFHPPRYGEVADG 245
    |||||

QY 241 TAVDIFSPGMCALEMVLEIQTNGDTRVTEEAIAARHSLSDPNRREFTLCLADPAR 300
    |||||
DB 246 TAVDIFSPGMCALEMVLEIQTNGDTRVTEEAIAARHSLSDPNRREFTLCLADPAR 305
    |||||

QY 301 PSAHSLFPH 309
    |||||
DB 306 PSVHSLFPH 314
    |||||

RESULT 3
US-08-933-750C-20
; Sequence 20, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hallman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
```

```
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ FILING DATE: September 23, 1997
/ CLASSIFICATION: 536
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0356 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 535 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: BRATUT03
/ CLONE: 864683
US-08-933-750C-20

Query Match          57.5%; Score 1534.5; DB 1; Length 535;
Best Local Similarity 61.6%; Pred. No. 7e-141;
Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6;

QY 17 EREDESESDIIESEPCGMRQREOVNQGNNPGLQSTFLAMDTEGVVNNELHFGD 76
    |||||
DB 44 EEBESEDSESLIESPCGMRQREOVNQNVRGIDSAVILAMDEBGEVNNVQVSE 103
    |||||

QY 77 RKAFAAHEKIQTFEQLVVDHPNIVKLHKYLDTSEACARYIFITEVSSGSLKQFLK 136
    |||||
DB 104 RKNYLOGEKRAVNDNIQLEHNIYVFHKYMDIKENKARVIFITEVSSGSLKQFLK 163
    |||||

QY 137 KTKKNHAKANARAKMCTQIISALSFLLHACSPPIIHGNTSDTITFIQHNGLIKIGSV 196
    |||||
DB 164 KTKKNHAKANARAKMCTQIISALSYLHSCDPPIIHGNTCDTIFIQHNGLIKIGSV-- 221
    |||||

QY 197 RIFSNALPPTALPDDLRSPIRARREELRNLFHPPRYGEVAD-CTAVDIFPGMCAL 255
    |||||
DB 222 -----APUTINNHVKTCEQKNLHFAPRYGEVNTVTTAVDIISFGMCAL 269
    |||||

QY 256 AVLEIQNGDTR-VTEEAIAARHSLSDPNRREFTLCLADPARPSAHSILFTRVLFE 314
    |||||
DB 270 AVLEIQNGESSYVQGEAISAQLLEDDPLQREFIQKLOGBPARPFAABELFPAALFE 329
    |||||

QY 315 VHSKLAAAGCFIQHYLMPENVVEEKTAMDLAVALR-PRRPLQMRYSVSEFME 373
    |||||
DB 330 VPSILTLAAHCTIVGQHWIPENALAEITKMDTSVALLEIPAGPBRBVQTLYSQSPAL 389
    |||||

QY 374 LDKFLIEDVRNGIYPLMNAATRPGLPRVLAPRPEV-----QKAKTPTEPPSPSENR 427
    |||||
DB 390 LDKFLIEDVRNGIYPLTA-----GLPRPQPOQEEVTSVVPVPSVKTPTEPPAVERRK 443
    |||||

QY 428 VIOMQCNLERSEDKARWHLTLVLVLEDRILHRLQTYDLIPTSDAODLASELVHYGFLHDD 487
    |||||
DB 444 VIOMQCNLESVEEGKHLTLILKLEDKLNHLSGDLMPNINIPBLAELVQLGISSEAD 503
    |||||

QY 488 RMKLAALFESTFLKY 502
    |||||
DB 504 OSRLTSLLEFTLNKF 518
    |||||
```



```

RESULT 4
US-09-234-613-20
Sequence 20, Application US/09234613
Patent No. 6132973
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guebler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/234,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750
FILING DATE: September 23, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 535 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITUT03
CLONE: 864683
US-09-234-613-20

57.5%; Score 1534.5; DB 2; Length 535;
Best Local Similarity 61.6%; Pred. No. 76-141;
Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6

17 EREDESDSILLESPCGRWQKRREQVNGNMPGLQSTFLAMDTEBGEVYVWNEIHFQD 76
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
44 EEESESESESEILLESPCGRWQKRREEVNQRVPIDSAIYLMDEBGEVYVWNEVQFSE 103
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
77 RKAPAAHEEIKQTVESQOLVVDHPRIIVLHKYMLTSEKAPVITTEVSSGSIKQPLK 136
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
104 RKNYVLQSEKRAVFDNLQLLEHLNIVFHKYMWADIKENKAVIITTEVSSGSIKQPLK 163
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 KTKQNHKMANARAWKRWCTQILSALSFYLHACSPPIIHGNLTSDTIFIOHGLIKGSVW 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
164 KTKQNHKTMNEKAWKRWCTQILSALSYLHSCDPIIHGNLTCDTIFIOHGLIKGSV-- 221
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
197 RIFSNALRPPLAPDLASPIRAEEELRNLFPPPEYGEVAD-GTAVDIFSGKCALDM 255
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db      222 -----APDTINNHYKTCBEGQNLHFFAPEYGEVNNVTAVDIYSPGCMALFM 269
Qy      226 AVLEIQTMGDR-VTEBAIARARHSLDPNMEFFLCLADPARPASHLSLFFRVLFE 314
Db      270 AVLEIQGEGESSYVQGEAISAIQLLEDPLOQREFLOKQCSPPARPTARELLFHPALFE 329
Qy      315 VHSKLTLAAHCFIOHYLMPENVVEBKTKAMDHLHVLAEPR-PRRPLQWRYSEVSFME 373
Db      330 VPSLKLTLAAHCFVGHQHMIPENALBEITKMDTSALHLEIPAGPGNEPVQTLYSQSPALB 389
Qy      374 LDKPLEDVNRNGIYPLNMPAATRPGLIPRYLAPRESEV-----OKAKTTPPEPDESTRK 427
Db      390 LDKPLEDVNRNGIYPLTAF-----GLPRPQOQGESEVISPVPVPSKTTTPREAEVETRK 443
Qy      428 VIOMQCNLERSEDKARMLTLLVLVEDRLHROLTYDLPTSDAODIASELVHYGFLHEDD 467
Db      444 VLMQCNLESYEEVGKHHLTLLKLEDKLNRHLSCDLMPENMIPETIAELVOLGFISEAD 503
Qy      488 RMKLAAPLESTFLKY 502
Db      504 QSRLLSTLEETLNKF 518

RESULT 5
US-09-949-016-7299
; Sequence 7299, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949.016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7299
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7299

Query Match      57.5%, Score 1534.5, DB 2, Length 562;
Best Local Similarity 61.6%, Pred. No. 7.5e-14;
Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6;

Qy      17 EREDESDSDILTESPCGRWOKRREONQGNMPLQSTFLAMDREGEVYVNNELHPRD 76
Db      71 EEESESEDESLTESPCGRWOKRREYVNOGRVPIIDAYLAMDREGEVYVNNVQFSB 130
Qy      77 RKAFAAHEKIQTFEBOLVLDHPYIVLYLKKYMLDTSACARVIFITEVSSGSLKQFLK 136
Db      131 RKYKLOGEKRAVAFNDLIQLEHNTIVYFHKYKMDIKENKARVITTEYMSGSLKQFLK 190
Qy      137 KTKKNHKAAMNARAKRWCTQIILSALSFTHACSPPIHGNLTSDTI FIOHNGLIKIGSVH 196
Db      191 KTKGNHKTMYNEKAKRWCTQIILSALSYHSCDPPIHGNLTCDTIFIQHNGLIKIGSV-- 248
Qy      197 RIFSNALRPPLALDDDLSPTRARREELRNHPPPEYGEVAD-GTAVDIPSGCMALFM 255
Db      249 -----ADDTINNHYKTCREBKXNLHFAPEYGEVNTAVDIYSPGCMALFM 296
Qy      256 AVLEIQTMGDR-VTEBAIARARHSLDPNMEFFLCLADPARPASHLSLFFRVLFE 314
Db      297 AVLEIQGEGESSYVQGEAISAIQLLEDPLOQREFLOKQCSPPARPTARELLFHPALFE 356
Qy      315 VHSKLTLAAHCFIOHYLMPENVVEBKTKAMDHLHVLAEPR-PRRPLQWRYSEVSFME 373

```

Db 357 VPSLKLAHACIVGHQHWI PENALBEITKQMDTSAYLAIEIPAGPGRBEPVQTLXSQSPALE 416
Qy 374 LDKFLEDVNRNGIYPLMNPATRPGLPRVLAPRPEV-----QKAKTTPRPFDSERK 427
Db 417 LDKFLEDVNRNGIYPLPAF-----GLPRPOQPOOEVTSPVVPSPVKTPRPEPAVEYERK 470
Qy 428 VIOMQCNLERSEDKARWH/TLVLVLEDR/LHROLTYDLLPTDSAODLASELYHGFLEBD 487
Db 471 VLMQCNIESVEBGVGH/HTLLKLEDKLNRHLSCDLMPNENIPELAAELVOLGFISEAD 530
Qy 488 RMKLAFLFESTFLKY 502
Db 531 QSRLLTSLEETLNKF 545

RESULT 6
US-09-023-942A-10
; Sequence 10, Application US/09023942A
; Patent No. 6479274
; GENERAL INFORMATION:
; APPLICANT: (US only) ANTALIS Toni Marie and HOOPER John David
; TITLE OF INVENTION: NOVEL MOLECULES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,942A
; FILING DATE: 13-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P05101/97
; FILING DATE: 13-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P04022/97
; FILING DATE: 18-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: International PCT Application
; FILING DATE: 13-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGIGLIO, FRANK S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 11168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742 4343
; TELEFAX: (516) 742 4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 535 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-023-942A-10
Query Match 56.5%; Score 1508.5; DB 2; Length 535;
Best Local Similarity 60.8%; Pred. No. 2.4e-138;
Matches 301; Conservative 60; Mismatches 105; Indels 29; Gaps 6;
Qy 17 EDESEDESDIIEESPCGRMQRQVNGNMPGLOSTFLAMDTEBEGVYVNNELHPGD 76
Db 44 EEESESESEIIEESPCGRMQRQVNGNMPGLOSTFLAMDTEBEGVYVNNELHPGD 103

Qy 77 RKAFAAHEEKIQTFEOLVLDHPNIVLKHKWLDTSEACARVITFEVSSGSLKQFLK 136
Db 104 RKNYLOGEKCAVFDNLIQLEHNTIVKFKYMDIKENKARVIFITYMSSGSLKQFLK 163
Qy 137 KTKGNKMANBARAKRWCTQILSALISFLHACSPPIIHGNTLSDTIFIOHNGLIKIGSVH 196
Db 164 KTKGNHQTMEKAWKRWCTQILSALISFLHACSPPIIHGNTLSDTIFIOHNGLIKIGSVH 221
Qy 197 RIFSAARPPALPDRLSPRARERELRNHAPRPEGEVAD-CTAVDISFGCALDM 255
Db 222 -----APDTINNHVKTCREQKHLHFAPBEYGEVNTVTAVDITSGFMCALGM 269
Qy 256 AVLEIQTNGDTR-VTEEAIPARARSLSDPNRREFILCCIADPARPARSAHSLFFRVLFE 314
Db 270 AVLEIQNGESSYVPOEAISSAIOLEBPLQREFIOKLOSEPAPRRPFRABELFTPALFE 329
Qy 315 VHSKLKLAHACFIOHQYIMPENVEEKTAKVDLHVAELPR-PRRPLQWRYSEVSFME 373
Db 330 VPSLKLAHACIVGHQHWI PENALBEITKQMDTSAYLAIEIPAGPGRBEPVQTLXSQSPALE 389
Qy 374 LDKFLEDVNRNGIYPLMNPATRPGLPRVLAPRPEV-----QKAKTTPRPFDSERK 427
Db 390 LDKFLEDVNRNGIYPLPAF-----GLPRPOQPOOEVTSPVVPSPVKTPRPEPAVEYERK 443
Qy 428 VIOMQCNLERSEDKARWH/TLVLVLEDR/LHROLTYDLLPTDSAODLASELYHGFLEBD 487
Db 444 VLMQCNIESVEBGVGH/HTLLKLEDKLNRHLSCDLMPNENIPELAAELVOLGFISEAD 503
Qy 488 RMKLAFLFESTFLKY 502
Db 504 QSRLLTSLEETLNKF 518

RESULT 7
US-09-345-473E-36
; Sequence 36, Application US/09345473E
; Patent No. 6558903
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin
; TITLE OF INVENTION: No. 6558903el kinases and Uses Thereof
; FILE REFERENCE: 35800/183781
; CURRENT APPLICATION NUMBER: US/09/345,473E
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 231
; TYPE: PRT
; ORGANISM: C. elegans
; US-09-345-473E-36

Query Match 20.3%; Score 543; DB 2; Length 231;
Best Local Similarity 56.3%; Pred. No. 1.2e-44;
Matches 98; Conservative 31; Mismatches 45; Indels 0; Gaps 0;
Qy 19 EDESEDESDIIEESPCGRMQRQVNGNMPGLOSTFLAMDTEBEGVYVNNELHPGRK 78
Db 23 DSDADGAEIIEESPCGRMQRQVNGNMPGLOSTFLAMDTEBEGVYVNNELHPGRK 82
Qy 79 AFAAHEEKIQTFEOLVLDHPNIVLKHKWLDTSEACARVITFEVSSGSLKQFLKKT 138
Db 83 NFAAHEEKINAVFDLTVLHNTLVKFKYMTDSKSKPRIFITFEVSSGSMANFLQRT 142
Qy 139 KKNHAKANARAKRWCTQILSALISFLHACSPPIIHGNTLSDTIFIOHNGLIKIG 192
Db 143 RKAGSLSLIKAKWKWTQILSALISFLHACSPPIIHGNTLSDTIFIOHNGLIKIG 196
RESULT 8
US-09-862-027-36
; Sequence 36, Application US/09862027
; Patent No. 6658418
; GENERAL INFORMATION:

```
/ APPLICANT: Hodge, Martin R.
/ TITLE OF INVENTION: No. 6558418el Kinases and Uses Thereof
/ FILE REFERENCE: 35800/234862
/ CURRENT APPLICATION NUMBER: US/09/862,027
/ CURRENT FILING DATE: 2001-05-21
/ PRIOR APPLICATION NUMBER: US 09/345,473
/ PRIOR FILING DATE: 1999-06-30
/ NUMBER OF SEQ ID NOS: 82
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 36
/ LENGTH: 231
/ TYPE: PRT
/ ORGANISM: C. elegans
US-09-862-027-36

Query Match      20.3%; Score 543; DB 2; Length 231;
Best Local Similarity 56.3%; Pred. No. 1.2e-44;
Matches 98; Conservative 31; Mismatches 45; Indels 0; Gaps 0;

QY 19 EDESEBDSILRESPCGRMQRREOVQGMPLQSLFPLMDPEBGEVYVNMELHRODRK 78
DB 23 DSDADGAEEILSESPDKRMSKRREOVQKQVDFGIDVAYLMDNETGEVYVNMVEQFSERK 82
QY 79 APAHSEKIQTVEEQLVLDHPNIVKLYKLYLDTSEACARVITTEVSSGSLKQFLKKT 138
DB 83 NFRAGSEKINAVVDNLQVLVHTLVKFKKMTDSKSEKPIITTEVSSGSMASFLQRT 142
QY 139 KKNHKMANAMARKWCTQIILSALSPFHACSPPIIHGNTLSDTIFIOHNGIKIG 192
DB 143 RKAGSSLSIKAMKMTQIILSALNYLHSSDPPIIHGNTLQNTVFIQNGIKIG 196

RESULT 9
US-09-345-473E-35
/ Sequence 35, Application US/09345473E
/ Patent No. 6558903
/ GENERAL INFORMATION:
/ APPLICANT: Hodge, Martin
/ TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
/ FILE REFERENCE: 35800/183781
/ CURRENT APPLICATION NUMBER: US/09/345,473E
/ CURRENT FILING DATE: 1999-06-30
/ NUMBER OF SEQ ID NOS: 62
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 35
/ LENGTH: 461
/ TYPE: PRT
/ ORGANISM: C. elegans
US-09-345-473E-35

Query Match      15.6%; Score 415.5; DB 2; Length 461;
Best Local Similarity 29.9%; Pred. No. 8.9e-32;
Matches 116; Conservative 57; Mismatches 92; Indels 123; Gaps 13;

QY 210 PDLRSPIRAEERELNLHFFPEYGEVADGT-AVDIFSGMCALEMAV---LEIQNGD 265
DB 14 PDINHNVKTCRNMKMYHIAPEYNNTELTSADIVYFGICSLIAVYIGLSGQNGS 73
QY 266 TR--VTEEAIAARASHLSDPNNKEFTLCCILADPARPSPASHLFFRVLFVYHSLKLLAA 323
DB 74 SSGPVTEDVIEKAIIRLEDPMODFIRQCLRKDPARPSARELLFQIILFEVHSLKLLSA 133
QY 324 HCEIHOYIMPENVEEKTAMDHLAVLAEPRPRRPQWRYSEVSF--MELDKFLEVDY 381
DB 134 HAIIVDSKTY--EDVSESARFIKD-NETIATSKLR---EMAYCQVAAPQVLEKFLDDY 186
QY 382 RAGIYPLAMFA-----ATRP-----LG 398
DB 187 RAGIYPLTAPLAPLADPSTTLTAYSTNPSSTLITTDISAPSSSTHPSANSTTAETSVNTS 246
QY 399 LPRVLAAP-----PEEVOK-----412
DB 247 LFGQSSQPSGTTTNTNGPSSIGKSASPEAVDKKIGVSTESTSKVEVEVNGANYTIGSS 306
```

```
QY 413 ----ATPPE-----PDSETRKYIQMOCNLEERSDRAKMLTLLVLEDRLRQ 459
DB 307 NGRDAGSPPEEERBNGERDMRLNRIHILEINVIENEB-----MSIVLLEDDQWHRQ 360
QY 460 LTYDLLPTSDAODLASELVHYGFLEDD 487
DB 361 LTTISINKDNPETLTENLTTHGPMQOLD 388

RESULT 10
US-09-862-027-35
/ Sequence 35, Application US/09862027
/ Patent No. 6558418
/ GENERAL INFORMATION:
/ APPLICANT: Hodge, Martin R.
/ TITLE OF INVENTION: No. 6558418el Kinases and Uses Thereof
/ FILE REFERENCE: 35800/234862
/ CURRENT APPLICATION NUMBER: US/09/862,027
/ CURRENT FILING DATE: 2001-05-21
/ PRIOR APPLICATION NUMBER: US 09/345,473
/ PRIOR FILING DATE: 1999-06-30
/ NUMBER OF SEQ ID NOS: 82
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 35
/ LENGTH: 461
/ TYPE: PRT
/ ORGANISM: C. elegans
US-09-862-027-35

Query Match      15.6%; Score 415.5; DB 2; Length 461;
Best Local Similarity 29.9%; Pred. No. 8.9e-32;
Matches 116; Conservative 57; Mismatches 92; Indels 123; Gaps 13;

QY 210 PDLRSPIRAEERELNLHFFPEYGEVADGT-AVDIFSGMCALEMAV---LEIQNGD 265
DB 14 PDINHNVKTCRNMKMYHIAPEYNNTELTSADIVYFGICSLIAVYIGLSGQNGS 73
QY 266 TR--VTEEAIAARASHLSDPNNKEFTLCCILADPARPSPASHLFFRVLFVYHSLKLLAA 323
DB 74 SSGPVTEDVIEKAIIRLEDPMODFIRQCLRKDPARPSARELLFQIILFEVHSLKLLSA 133
QY 324 HCEIHOYIMPENVEEKTAMDHLAVLAEPRPRRPQWRYSEVSF--MELDKFLEVDY 381
DB 134 HAIIVDSKTY--EDVSESARFIKD-NETIATSKLR---EMAYCQVAAPQVLEKFLDDY 186
QY 382 RAGIYPLAMFA-----ATRP-----LG 398
DB 187 RAGIYPLTAPLAPLADPSTTLTAYSTNPSSTLITTDISAPSSSTHPSANSTTAETSVNTS 246
QY 399 LPRVLAAP-----PEEVOK-----412
DB 247 LFGQSSQPSGTTTNTNGPSSIGKSASPEAVDKKIGVSTESTSKVEVEVNGANYTIGSS 306

QY 413 ----AKTPE-----PDSERKVIQMCNLEERSDRAKMLTLLVLEDRLRQ 459
DB 307 NGRDAGSPPEEERBNGERDMRLNRIHILEINVIENEB-----MSIVLLEDDQWHRQ 360
QY 460 LTYDLLPTSDAODLASELVHYGFLEDD 487
DB 361 LTTISINKDNPETLTENLTTHGPMQOLD 388

RESULT 11
US-09-345-473E-34
/ Sequence 34, Application US/09345473E
/ Patent No. 6558903
/ GENERAL INFORMATION:
/ APPLICANT: Hodge, Martin
/ TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
/ FILE REFERENCE: 35800/183781
/ CURRENT APPLICATION NUMBER: US/09/345,473E
/ CURRENT FILING DATE: 1999-06-30
```



```

Db      319  --CFSRIPAGYPIREIIRDCIRVRERBSTVKQL-----LVDDFPTPEDLIGIRVEIKN 371
Qy      343  KAMDIAVLAELPRPRPPQMKRYSEVSEFMELD--KLELVKRGCIYLMNPAATRPLGLP 400
Db      372  RDADLNDLNEIOMQLVYDEKRRKQYRFKENBGLQAPFIEN-----414
Qy      401  RVLAAPPEEYQAKATPPEPDSRTYVIOQCNLSESEDKARHMLTLVLVEDRLHROL 460
Db      415  ----DSDEVVQCMIEQOHIPDETRNITKL-----IKDKV-----DAFRDR 453
Qy      461  TYDLLPTDSAODLASELVHYGFLHEDDRMLAA 493
Db      454  DHRLLEIKRAKEEBERIREAEIKELRLABA 486

RESULT 14
US-09-862-027-40
; Sequence 40, Application US/09862027
; Patent No. 6858418
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; FILE OF INVENTION: No. 6858418el Kinases and Uses Thereof
; TITLE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 1601
; TYPE: PRF
; ORGANISM: C. elegans
US-09-862-027-40

```

Query Match 13.8%; Score 368.5; DB 2; Length 1601;

Best Local Similarity 25.0%; Pred. No. 2,2e-26; Matches 128; Conservative 88; Mismatches 188; Indels 109; Gaps 20;

```

Qy      10  RAREREREDS-----EDESIDI--LESPPGMRQKRREOVNQMPLQSTPL 57
Db      54  RAKEQEKREBEKKAARRIDVEDDPAQEKIDKSKNGRFLKFDDELGRS---FKTVPR 110
Qy      58  AMDTGEGVYVNNELHFGDKAPAAHEKIQTVFEQVLVDHNYIKLHKWLDTSACA 117
Db      111  GLDTEGVAVAWCELO--ESKLNKTERQRFREAEMLKLOHPNIYRFYDYW--ESADLCG 167
Qy      118  R---VIFITEYVSSGSLKQFLKTKTKK-NHKAMNABAKRWCTOILSALSFHACSPPIIH 173
Db      168  KRYIYLVLELMTSGTLKMYLKRKRINIKV---LKSVCROILKGLSFLHTRNPPVILH 223
Qy      174  GNLTSPTFIQ--HNGLIKIGSVNHRIFSNALRPPTALPDDLRSPIAREBELNLHFFPP 232
Db      224  RDLKCNIFITGTTGSAVKIGDLGLATLKNK-----SPAKSVIGTPPEFMAP 268
Qy      233  EYGVADGTAVDIFSFGMCLEMAVLEI---QTNQDTRVTEEAIAARHSLSDPNKREPI 289
Db      269  EWEYEMVDSVDVAYAGMCLLEMVTGVPYSECMNPATYIRKVISGVKPE-----318
Qy      290  LCCLADPARPARSAHSILPH--RVLPVNS-LKLLAHCFIQHQLMPENV-----VEEKT 342
Db      319  ---CFSRIPAGYPIREIIRDCIRVRERBSTVKQL-----LVDDFPTPEDLIGIRVEIKN 371
Qy      343  KAMDIAVLAELPRPRPPQMKRYSEVSEFMELD--KLELVKRGCIYLMNPAATRPLGLP 400
Db      372  RDADLNDLNEIOMQLVYDEKRRKQYRFKENBGLQAPFIEN-----414
Qy      401  RVLAAPPEEYQAKATPPEPDSRTYVIOQCNLSESEDKARHMLTLVLVEDRLHROL 460
Db      415  ----DSDEVVQCMIEQOHIPDETRNITKL-----IKDKV-----DAFRDR 453
Qy      461  TYDLLPTDSAODLASELVHYGFLHEDDRMLAA 493

```

Db 454 DHRLLEIKRAKEEBERIREAEIKELRLABA 486

RESULT 15

US-09-270-767-31733

; Sequence 31733, Application US/09270767

; Patent No. 6703493

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 31733

; LENGTH: 110

; TYPE: PRF

; ORGANISM: Drosophila melanogaster

US-09-270-767-31733

```

Query Match 13.7%; Score 364.5; DB 2; Length 110;
Best Local Similarity 61.0%; Pred. No. 1e-27;
Matches 72; Conservative 14; Mismatches 21; Indels 11; Gaps 2;

```

```

Qy      118  RVIFITEYVSSGSLKQFLKTKTKK-NHKAMNABAKRWCTOILSALSFHACSPPIIHGNT 177
Db      3  RVIFITEYVSSGSLKQFLKTKTKK-NHKAMNABAKRWCTOILSALSFHACSPPIIHGNT 62
Qy      178  SPTFIQNHGLIKIGSVNHRIFSNALRPPTALPDDLRSPIAREBELNLHFF--PREY 234
Db      63  CDSIFIQNHGLVKGISV-----VPDAVYSVRGRERERERGAHYFOADEY 110

```

Search completed: January 12, 2006, 11:23:01

Job time : 26 secs

THIS PAGE BLANK (USPTO)


```

Db 209 -----EFIOKCLQSEPARPTARELLFHPALFEVP 238
Qy 317 SLKLLAHCIOHOYIMPENVEKTKAMDHLAVLAELRP-PRRPLQOYRYSVSMELD 375
Db 239 SLKLLAHCIOHOYIMPENVEKTKAMDHLAVLAELRP-PRRPLQOYRYSVSMELD 286
Qy 376 KFLVDVNGIYPIPLMNAATRLCLPRVLAAPREEV-----QKAKTPTEPPDSETRKYI 429
Db 287 -----TLNGIYPLTA-----GLPRHQOQGEVSVSPVPSVKTPTEPPAVENTRYKV 335
Qy 430 QMOCNTERSEDKARMLTLTLVLEDRHLROLYVDLPTSDASOVLASELYHGFLE 485
Db 336 LMOCNTERSEVGVKHHLLTLKLEDKLNRHLSCLMPNENIPETAELVQLGFISE 391

```

```

RESULT 2
US-60-752-355-38255
; Sequence 38255, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 38255
; LENGTH: 554
; TYPE: PRF
; ORGANISM: Oryza sativa
US-60-752-355-38255

```

Query Match 7.5%; Score 199; DB 8; Length 554;

Best Local Similarity 22.8%; Pred. No. 1.1e-10; Matches 115; Conservative 68; Mismatches 167; Indels 154; Gaps 25;

```

Qy 2 AAEPRAREREREREDESEDSILESPPC-----GRWQKR-----EQVNOG 47
Db 125 SSBHPLRP-----PSSPCSSRRVVSQKKKLGSGTFGGVYQG 165
Qy 48 -NMPGLQSTFLAMDTBEGVEVWNEHFGDKAPAAHEKIQTVPEQLV---VDHPNIV 103
Db 166 FNSGCG-----MCAKEVVISD-----SNSKECLRLHOEIVLSLSHPNIV 211
Qy 104 KLMHWLDTSEACARVIFITEYSSGSLKQFLKTKKNHKAAMARAKMCTQILSALSP 163
Db 212 QY--YSGDLSSETLSYL--EYVSGSINKLQE---YAFGEAVLRNTTAQILSGLAY 263
Qy 164 LHAESPPIHGNLTSDTIFIOHNGLIKI--GSVWHRIFSNALRPPTALPDDLRSPIRAE 220
Db 264 LHGRN--TVHRDIKGANILVDPNGDIKLADFGMAKH-----ISAHTSIKSPKSGPYWMA 315
Qy 221 REELRLHFFRPPEYGEVADSTAVDIFSFGCALEMAVLE---IQNGDTRVTEBAIARAR 277
Db 316 PEVIYNTNGY-----SLSYDWSIGCTIEMTARPPWIOYEG-----VAALFKIG 361
Qy 278 HSLSDPNM-----REFILCCLARDPARPSASHLFFRVLFV----- 315
Db 362 NSKOITPDIDPHLSFEAKNPLKLCIQRDPAKPTPAQOLMEHPYKDLVANKSFGSGMTRDT 421
Qy 316 -----HSLKLLAHCIOHOYIMPENVV-----EETKAMDLHVLAELEP 355
Db 422 FPTSFGKAMVQTSNRSLSPRLDPDLTMNLOVPTSAISISTRISAINPNVMMNS 481
Qy 356 RPRRP---PL-QMRYSEVSPMELDKF-----LEDVNGIYPLMFAATPPLGLPRV 402
Db 482 LPVSPCSSPLRQYRQSNRSRCPSPRPHAYAGANYSINNALYP-----KRP---SNH 532
Qy 403 LARPEEVOKAKTPTEPPDSETR 426
Db 533 VDPPELTSIQKAP---PFDSPRR 553

```

RESULT 3

```

US-60-752-355-5150
; Sequence 5150, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 5150
; LENGTH: 833
; TYPE: PRF
; ORGANISM: Oryza sativa

```

```

; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (695)..(695)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (726)..(726)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-60-752-355-5150

```

Query Match 7.3%; Score 196; DB 8; Length 833;

Best Local Similarity 25.3%; Pred. No. 4e-10; Matches 83; Conservative 61; Mismatches 112; Indels 72; Gaps 17;

```

Qy 32 SPGSMQKREOVNOGNMFGLOSTF---LAMDTBEGVEVWNE--LHFGDKAPAAHEE 85
Db 404 SPGSRWK-----GKLIG-RGTGHHVYVGFNSDGEKAMKEVTLFLDDPKS---XE 451
Qy 86 KIQTVPEQLV---VDHPNIVLHKWMDTSACARVIFI-TEYVSSGSLKQFLKTKKN 141
Db 452 SAKQLQGEISLSRLQHNPVIQY-----GSETVDKITYLETVYSGSINKLQE--- 502
Qy 142 HKAMNARAKMCTQILSALSPFIHGNLTSDTIFIOHNGLIKISGVWHRIFSN 201
Db 503 YQGLGEQALRSTYQQLISGLATYLHAKN--TVHRDIKGANILVDPGSRVTLADFGMAKHIN 560
Qy 202 ALRPTALPDDLRSPIRAREELRLHFFRPPEYGEVADG--TAVDIFSFGCALEMAVLE 259
Db 561 GQCCFPSPKSG--SP-----YWWAPRVIKNSNCNLAVDIWSIGCTVLEWATSK 606
Qy 260 ---IQNGDTRV---TERAIRARHSLSDPMRREFILCCLARDPARPSASHLFFRV 312
Db 607 PMSQYEGIAAMFKIGNSKELRPIPDHLSBPQ-KDIFRKLQRPDSQRPAMELL----- 660
Qy 313 FEVHSLKLLAHCIOHOYIMPENVVEE 340
Db 661 -----QHPVQKAVSLKSVLSE 678

```

RESULT 4

```

US-60-752-355-13134
; Sequence 13134, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 13134
; LENGTH: 894
; TYPE: PRF

```

```

; ORGANISM: Oryza sativa (japonica cultivar-group)
US-60-752-355-13134

```

Query Match 7.3%; Score 196; DB 8; Length 894;

Best Local Similarity 25.3%; Pred. No. 4.5e-10;

Matches	83;	Conservative	61;	Mismatches	112;	Indels	72;	Gaps	17
Qy	32	SPCGMWQRREOVGNMNGLOSTP----	LANDTEEGEVVWNE--LHFGDKAPAAHEE	85					
Db	404	SPGSRMVK-----GKLIG-RGTGHHVYVGNSSGEMCAKEVTLFLDDPKS----	KE	451					
Qy	86	KIQTVFEQVLV---VDHNIVYKHKYMLDTSACARVFI--TEVYSSGSLKQFLKTKKN	141						
Db	452	SAKQGGQGISLSLRQHFNIVQY-----GSETVDKLYLEYEVSGSINHLQDE----	502						
Qy	142	HKAMNARAKWKCOTQIISLASFHACSPPIIHGNLTSDTPIQHNGLIKIGSVHRIFSN	201						
Db	503	YGOJGEQAIRSYTQOILISGLAVLHAKN--TVHRIKGNILVDPGRVYKLDFFGMAKHIN	560						
Qy	202	ALRPETALPDDLRSPRIABREBELRNLFHFFPEYGEVADG--TAVDIFSGFCALEMAVLE	259						
Db	561	GOQCFPSFKG--SP-----YMAPEVIRKNSGNCNLAVDIWSIGCTVLEMAISK	606						
Qy	260	---IQTNGDTRY---TEEAJARARHSLSDPMMREFIICCLARDPARPSASHLFFHVL	312						
Db	607	PPMSQYEGIAAFKIGNSKELEPPIPDHISEPG-KDIFIRKICLQRPDSQRTAMELL----	660						
Qy	313	FEVHSLKLLAAHCFIQHOYLMPENVYEE	340						
Db	661	-----QHPIVQKAVSLSEKSTLSE	678						
RESULT 5									
	US-60-752-355-11780								
	Sequence 31780, Application US/60752355								
	GENERAL INFORMATION:								
	APPLICANT: Abad, Mark et al.								
	TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits								
	FILE REFERENCE: 38-21(53720)B								
	CURRENT APPLICATION NUMBER: US/60/752,355								
	CURRENT FILING DATE: 2005-12-21								
	NUMBER OF SEQ ID NOS: 52803								
	SOFTWARE: PatentIn version 3.3								
	SEQ ID NO 31780								
	LENGTH: 894								
	TYPE: PRT								
	ORGANISM: Oryza sativa								
	FEATURE:								
	NAME/KEY: misc_feature								
	LOCATION: (285)..(289)								
	OTHER INFORMATION: Xaa can be any naturally occurring amino acid								
	US-60-752-355-11780								
Query Match		7.3%	Score 196;	DB 8;	Length 894;				
	Beet Local Similarity	25.3%;	Pred. No. 4,5e-10;						
Matches	83;	Conservative	61;	Mismatches	112;	Indels	72;	Gaps	17
Qy	32	SPCGMWQRREOVGNMNGLOSTP----	LANDTEEGEVVWNE--LHFGDKAPAAHEE	85					
Db	404	SPGSRMVK-----GKLIG-RGTGHHVYVGNSSGEMCAKEVTLFLDDPKS----	KE	451					
Qy	86	KIQTVFEQVLV---VDHNIVYKHKYMLDTSACARVFI--TEVYSSGSLKQFLKTKKN	141						
Db	452	SAKQGGQGISLSLRQHFNIVQY-----GSETVDKLYLEYEVSGSINHLQDE----	502						
Qy	142	HKAMNARAKWKCOTQIISLASFHACSPPIIHGNLTSDTPIQHNGLIKIGSVHRIFSN	201						
Db	503	YGOJGEQAIRSYTQOILISGLAVLHAKN--TVHRIKGNILVDPGRVYKLDFFGMAKHIN	560						
Qy	202	ALRPETALPDDLRSPRIABREBELRNLFHFFPEYGEVADG--TAVDIFSGFCALEMAVLE	259						
Db	561	GOQCFPSFKG--SP-----YMAPEVIRKNSGNCNLAVDIWSIGCTVLEMAISK	606						
Qy	260	---IQTNGDTRY---TEEAJARARHSLSDPMMREFIICCLARDPARPSASHLFFHVL	312						
Db	607	PPMSQYEGIAAFKIGNSKELEPPIPDHISEPG-KDIFIRKICLQRPDSQRTAMELL----	660						
Qy	313	FEVHSLKLLAAHCFIQHOYLMPENVYEE	340						
Db	661	-----QHPIVQKAVSLSEKSTLSE	678						

```

Db      661 -----OHPFVQKAVSLKSVLSE 678

RESULT 6
US-60-752-355-11788
; Sequence 11788, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 11788
; LENGTH: 600
; TYPE: PRN
; ORGANISM: Glycine max
US-60-752-355-11788

Query Match      7.3%; Score 195.5; DB 8; Length 600;
Best Local Similarity 23.7%; Pred. No. 2,86-10;
Matches 101; Conservative 62; Mismatches 151; Indels 113; Gaps 20;

QY      32 SPCGRKQGRREQVQGNMPLQSTF---LAMDTEG-----VEYVMELHFQDKXAF 80
Db      189 SNVSKMKR-----GKLLG-RGTFGHVYVIGFNSENGOMCAIKKVKVSD----- 231
QY      81 AAHEKIQTVTEQOLVLD---HPIIVLHKYKWDTESEACARVFI-TEYVSSGLKQFTK 136
Db      232 QTSRECKLQMLQEIINLQUSHPIVOYH---GSELVEBSLSVLEYVSGSIIKLLQ 286
QY      137 KTKKNHKAHMAHAKRKCCTOILSLSEFLACSPPIIHGNLSDTIFIQHGLIKI---GS 193
Db      287 E---YSGFKEPVYQNTROQVSLAVLHGN---TVHRDIGANILVDPNGEITLADPGM 340
QY      194 VWHRIFSNALRPPTALPDDLRSPIRAEREELRNLFHPPEYGEVADGTAVDIFSFGMAL 253
Db      341 AKH-----INSSASMLSPKSGPFWMAPEVVMNTNGY-----SLPVDIWSLGCTII 385
QY      254 EMALVLEIQTDNPTVTEBEATARAHSLSDPMRE-----FTLCLAPDAPRPSAHS 305
Db      386 EAAVSKPPMQYEGVA--ALFKIGNSKDMPEIPEHLSNDANKFTKLCIQRDPLARPTAHK 443
QY      306 LIFRVLFEVHSLKTLAAHCFIQHQYLMPEVVEEKTAKMDLHVAELP---RPRRP 361
Db      444 LL-----DHPRIQK-----SATKANVSVITRAFPMPDGSKTRP 479
QY      362 -LQMRVSEVSEFMELDKLELDVRNGIYPLMFAATRPGLPVLAPPEEVQAKATPTPEP 420
Db      480 VLESNSRSTISITSDG-----DFASKPLAABRALRSPDNTMTISLPSVP 526
QY      421 FDSSTRK 427
Db      527 SSSPLRR 533

RESULT 7
US-60-752-355-1875
; Sequence 1875, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1875
; LENGTH: 611
; TYPE: PRN
; ORGANISM: Nicotiana benthamiana

```

US-60-752-355-1875

Query Match 7.3%; Score 195.5; DB 8; Length 611;
 Best Local Similarity 25.0%; Pred. No. 2.8e-10;
 Matches 108; Conservative 57; Mismatches 134; Indels 133; Gaps 24;

QY 36 RMOKREOVNOGMPGLQSTF-----LAMDTEG-----VEVNMELHFGDRKAPAAIE 84
 DB 206 KMKK-----GKLLG-RGTFGVVYLGFRNQGMAIKEVRVSD-----QTKK 248
 QY 85 EKIQTVEQLV---VDHPNIVKLHAKWYLDTSEACARVIFTEYSSGSLKQFLKTKTKN 141
 DB 249 ECKJLNQBIILSNLSHPNIVAYYGSBLDDETLVYL---EYVSGSIIHKLQE----- 300
 QY 142 HKAMBARAKMCTOTLSALSEFLHACSPTIIGHNLTSDTIFIOHNLIXI---GSVWHRI 198
 DB 301 YGAFREPVYQNYRQILSLGSLFHAAN--TVHNDIGANLVDPNEIKLADFGMAKH-I 357
 QY 199 FSNALRPPTLPDDLSPIRAREEDLNHLFPPEYGEVADGPAVDIFSGMCALEMA-- 256
 DB 358 TSSSL-----VLSFKSPYMAPEVVMNTSGY-----GLPVDIWSLGCALILEMASS 403
 QY 257 -----VLEIQNTGTRVTEEARARHSLSDPMMEFLICCLADPARPSASHLL 307
 DB 404 KPFWOYEGVAAIFKIGNSKDFPEI---PDHLSND--AKKFILCLQREPSARPTASQLL 458
 QY 308 FHRVLFVHSLKLLAHCFIOHQYLMPENVSEKTKAMDLHATLAPR---RRPQ 363
 DB 459 -----EHPFVKQ-----STTVTHGVTKAYPPSPDGNRTPPV- 493
 QY 364 WRXSEVSMELDKLEVDNRNGIYPLM-NFPA-----TPPLGLPVLAPPEVECAKATP 416
 DB 494 -----LDSGNISPTKGNVASHPVITISRLSCPR-----EIVTITIS 532
 QY 417 TP-EPPDSETRK 427
 DB 533 LPVSPTSPLRQ 544

RESULT 8

US-60-752-355-42688
 ; Sequence 42688, Application US/60752355
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark et al.
 ; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
 ; FILE REFERENCE: 38-21(53720)B
 ; CURRENT APPLICATION NUMBER: US/60/752,355
 ; CURRENT FILING DATE: 2005-12-21
 ; NUMBER OF SEQ ID NOS: 52803
 ; SOFTWARE: Patent in version 3.3
 ; SEQ ID NO 42688
 ; LENGTH: 567
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 US-60-752-355-42688

Query Match 7.2%; Score 193.5; DB 8; Length 567;

Best Local Similarity 25.3%; Pred. No. 3.9e-10;
 Matches 86; Conservative 52; Mismatches 107; Indels 95; Gaps 17;

QY 2 AABEPAPRAREREREREDESESDILESPPC-----GRWOKR-----EQVNOG 47
 DB 203 SSSHPLEPR-----PGSPSSSRVSSQMKKGLLQSGTFGVYQG 243
 QY 48 -NMPGLQSTFLANDTEGVEVNMELHFGDRKAPAAHEKIQTFVQLV---VDHPNIV 103
 DB 244 FNSGCG-----MCAIKVAVISDD-----SNSKECLRQHQBIYVLSQLSHPNIV 289
 QY 104 KHKWYLDTSEACARVIFTEYSSGSLKQFLKTKTKNKKAMBARAKMCTOTLSALSP 163
 DB 290 QY--YSSDLSSETLSVYL---EYVSGSIIHKLQE---YAGFBAVLRNTAQTILGLAY 341
 QY 164 LHACSPPIIGHNLTSDTIFIOHNLIXI---GSVWHRIFSNALRPPTLPDDLRSPIRAE 220

DB 342 LHGRN--TVHNDIKANILVDPNGDILKADFGMAKH-----ISATSIKSPKSGPYMA 393

QY 221 REELNHLHFPPEYGEVADGPAVDIFSGMCALEMAULE--IQNTGTRVTEEARAR 277

DB 394 PEVIMNTNGY-----SLSDIWSLGCCTIIEAATRAPPIVQYEG-----VAAIFKIG 439

QY 278 HSLDDPM-----REFILCLADPARPSASHLLFH 309

DB 440 NSKIDPDIPLHLSFAKNFLKCLQRPAPARTPAQLMEH 479

RESULT 9

US-60-752-355-35913
 ; Sequence 35913, Application US/60752355
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark et al.
 ; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
 ; FILE REFERENCE: 38-21(53720)B
 ; CURRENT APPLICATION NUMBER: US/60/752,355
 ; CURRENT FILING DATE: 2005-12-21
 ; NUMBER OF SEQ ID NOS: 52803
 ; SOFTWARE: Patent in version 3.3
 ; SEQ ID NO 35913
 ; LENGTH: 1484
 ; TYPE: PRT
 ; ORGANISM: Candida albicans SC5314
 US-60-752-355-35913

Query Match 7.2%; Score 193; DB 8; Length 1484;

Best Local Similarity 22.6%; Pred. No. 1.9e-09;
 Matches 79; Conservative 56; Mismatches 154; Indels 60; Gaps 13;

QY 10 RAREEREREDESESDILESPPC-----RMOKREOVNOGMPGLQ 53
 DB 1135 KIEERRRMLQEBFQSIGRLVDVSDSEYQFLTLASSFSSVSIRWOK-GACIGRT--FG 1190
 QY 54 STEFLAMDTGEGVAVNMELHFGDRKAPAAHEKIQTFVQLVDPNPI-----VCLHKY 108
 DB 1191 QVFSVAVNDTGGVMAVKETTFHDQSQVKTIVSICEKMTVLMNLHNPVYQFGVEVHR- 1249
 QY 109 WLDTEACARVIFTEYSSGSLKQFLKTKTKNKKAMBARAKMCTOTLSALSPHACS 168
 DB 1250 -----DKYIIFMEFCGSLAGLL-----THGRIDEMVYQVVALQWLEGLAYLHQSG 1297
 QY 169 PRITGNLTSDTIFIOHNLIXI---IGSVWHRIFSNALRPPTLPDDLRSPIRAEKEIR 225
 DB 1298 --VHRDIPENVLLDHNGVIXYVDFGA-KVIASGRITIGMTNSSLKSVKRDGHNL 1354
 QY 226 NLHFPPEY--GEVADGTA-----VDIPSGMCALEMAULE--IQNTGTRVTEEAR 274
 DB 1355 NSMTGTPMTSPBAITGSTDSDSGVVDIWSLGCCTIIEAATRAPPIVQYEG-----VAAIFKIG 1414
 QY 275 RARHSLSDPM-----REFILCLADPARPSASHLLFHRVLFVHSL 318
 DB 1415 GHKPOLPSPBQSESPQRTFLARCLEHDPYKRSAYVELADPMWVEIRHM 1463

RESULT 10

US-60-752-355-24004
 ; Sequence 24004, Application US/60752355
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark et al.
 ; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
 ; FILE REFERENCE: 38-21(53720)B
 ; CURRENT APPLICATION NUMBER: US/60/752,355
 ; CURRENT FILING DATE: 2005-12-21
 ; NUMBER OF SEQ ID NOS: 52803
 ; SOFTWARE: Patent in version 3.3
 ; SEQ ID NO 24004
 ; LENGTH: 786
 ; TYPE: PRT
 ; ORGANISM: Zea mays

US-60-752-355-24004

Query Match 7.2% Score 191.5; DB 8; Length 786;

Best Local Similarity 21.5%; Pred. No. 1e-09; Indels 161; Gaps 26;

Matches 109; Conservative 78; Mismatches 160; Indels 161; Gaps 26;

QY 2 AAEPAARRARREREREDESEDESDILESPC-----GRWQKREOVQGNMPC---LQ 53

DB 356 SSQPLP-----LPPGSPCLSSSLQMK-----GKLGSSTPG 369

QY 54 STLAMDTEG-----VEVWNEHLFGGRKAPAAHEEKIQVFEQVL---VDHENV 103

DB 390 QVWGFSEGGOMCAIKKVVISD-----SNKESLRQLNGRIVLLSQSHPNIV 440

QY 104 KLRKYMWDTEACARVFI--TEYVSSGSLQKFLKTKKNNKAMARAKMCKCTQIISLS 162

DB 441 QYV-----GSDLCNETSVLEEVYSGS IHLQOE---YGPGEAVLRNYTAQIISGLA 491

QY 163 FLHACSPPIHGNLTSPTFIQHNGLIKI---GSVWHIFSNALRPPTALPDLRSPIRA 219

DB 492 YLHGRN--TVHRDIKANNILVDPNGDIKLADFGMAGH-----ISATYISKFKSGSPYMW 543

QY 220 EREELRNHLFPPEYGEVADGTAVDIFSGMCALEMAVLE---IQNGDTRVTEBAIARA 276

DB 544 APEVIMNSNGY-----SLSDIWSLCTILEMATKAPWMSQYEGVAIAFK--IGNS 592

QY 277 RHSLDPPN-----MREPIICCLARDPARRPSAHSILFH----- 309

DB 593 KQIPDIPNNLSBAKGLKLCLODPAPAPTAQMLDHPVKQATVRSRSSITRDMFP 652

QY 310 -----RYLFEVHSLKLAACHFIOHQYLMPEV-----VEEKTAMDIAVL 351

DB 653 NSTDGKNSRKVITSSVRSLSR---LRDPILGRNLPGPSPISPSSRIALNLSNVR 709

QY 352 AELPRRRP---PL-QMRSEVSFMELDKF-----LEDVRNGIYPLMNAATRPPLG 398

DB 710 MNMSLVSFSSPLRQYRQSNRCLNSPPHAYSAGANYPINNMLYF-----TRP-- 761

QY 399 LPRVLAAPPEEVOKATPTPEPDSER 426

DB 762 -SSGLTDPWLEISQVKTQT---FDSPPR 785

RESULT 11

US-60-752-355-10912

Sequence 10912, Application US/60752355

GENERAL INFORMATION:

APPLICANT: Abad, Mark et al.

TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits

FILE REFERENCE: 38-21(53720)B

CURRENT APPLICATION NUMBER: US/60/752,355

CURRENT FILING DATE: 2005-12-21

NUMBER OF SEQ ID NOS: 52803

SOFTWARE: Patentin version 3.3

SEQ ID NO 10912

LENGTH: 627

TYPE: PRT

ORGANISM: Lotus corniculatus var. japonicus

US-60-752-355-10912

Query Match 7.1% Score 190.5; DB 8; Length 627;

Best Local Similarity 24.9%; Pred. No. 8.8e-10; Indels 101; Gaps 21;

Matches 106; Conservative 57; Mismatches 162; Indels 101; Gaps 21;

QY 28 ILEESPC--GRWQKREOVQGNMPCIGSTF---LAMDTEGVEVWNEHL--FGDRKAF 80

DB 210 VLENNCTNLSKMKK-----GKLGG-RGTFGHVYLGPSNENGOMCAIKKVFSDK-- 259

QY 81 AAHEEKIQVFEQVLVD---HPNIVKLAHKYMLDTEACARVFIETEVSSGSLKQPLKK 137

DB 260 -TSKEKLCQNLQNSHNPVVOY--YGELEESLSVYL--EYVSGS IHLQOE 314

QY 138 TKNHKAARAKRWCTQILSLSLHACSPPIHGNLTSPTFIQHNGLIKISVWHR 197

DB 315 ----YGAKEPVQNTROQVSGLAYIH--SRATYVRDIKGANILVDPNKEITLADPG-- 366

QY 198 IFSNALRPPTALPDLRSPIRAREBELRNHLFPPEYGEVADGTAVDIFSGMCALEMA- 256

DB 367 -MSKHINSAASMLSPFGSKPYWMAPEVVMNNGY-----GLPVDISLGGCTILEMAT 416

QY 257 -----VLEIQNGDTRVTEBAIARASHLSDPNMRKFLICCLADPARRPSAHS 306

DB 417 SKPPWSQFEGVAALIFKIGSKMPB---IPEHLSD--DAKNFTKQCLQRDPLARPTAQSL 471

QY 307 LFRVLFEVHSLKLAACHFIOHQYLMPEVVEEKTAMDIAVLALP---RPRRPL 362

DB 472 L-----NHPPIRDQ-----SATKVNASTITRDAFPMSGSRTPV 507

QY 363 QMRSEVSFMELDKFLEDVRNGIYPLMNAATRP--LGPLRVLAAPPEEVOKATPTPEPF 421

DB 508 LEPHSN-----RSSITLDDVATKPLAAVRTLNPRDSTRITISLPVSPS 554

QY 422 DSETRK 427

DB 555 SSPLRQ 560

RESULT 12

US-60-752-355-30838

Sequence 30838, Application US/60752355

GENERAL INFORMATION:

APPLICANT: Abad, Mark et al.

TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits

FILE REFERENCE: 38-21(53720)B

CURRENT APPLICATION NUMBER: US/60/752,355

CURRENT FILING DATE: 2005-12-21

NUMBER OF SEQ ID NOS: 52803

SOFTWARE: Patentin version 3.3

SEQ ID NO 30838

LENGTH: 555

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-60-752-355-30838

Query Match 7.0% Score 186.5; DB 8; Length 555;

Best Local Similarity 23.3%; Pred. No. 1.8e-09; Indels 101; Gaps 20;

Matches 97; Conservative 55; Mismatches 164; Indels 101; Gaps 20;

QY 42 EQVQGNMPCIGSTFPLAMDTEGVEVWNEHLF---GDRKAPAAHEEKIQVFEQVLVD 98

DB 8 EQGKGS---FQSLALVRHKQEKKKYVLKKIRLARGSDARSAHQE-----MELISTVR 59

QY 99 HPNIVKLAHKYMLDTEACARVFIETEVSSGSLKQFLKTKKNNKAMARAKMCKCTQIL 158

DB 60 NPEVVEYKDSWVE--KGC--YVCIVIGYCOGDMTDTIKRACGVH--FPEKLCQWMLVQL 114

QY 159 SALSFLHACSPPIHGNLTSPTFIQHNGLIKISVWNEHLIFSNALRPPTALPDLRSPIR 218

DB 115 MALDYTH--SNHILHADVCSNIFLTKEODIRLD-----FGLA--KLITSDDLSSVAV 164

QY 219 AEREELRNHLFPPEYGEVADGTAVDIFSGMCALEMA-----VLEIQNGDTRVTEB 271

DB 165 GTPS-----YMCPELADIPYGSKSDIWSLGCCTEMAAHKPPFKASDVQT----- 210

QY 272 AIAARASHLSDP-----MREPIICCLARDPARRPSAHSILFHRVLEEVHSLKLAH 324

DB 211 LITTKHKLIMDIPAMVSGSFRGLIKSMILRKPELRPSANEYL----- 253

QY 325 CFIOHQYLMPEVVEEKTAMDIAVLALP---PLQMRSEVSFMELDKFLEDVR 382

DB 254 ---NHPILQPY-----ISMVYKLESBSPRSTPFLQ-----PSEBDATLKERR 292

QY 383 NGIYPLMNAATRPPLGLPRVLAAPPEEVOKATPTPEPDSERTRKTVQMOCNLSE 439

DB 293 RS-----SFSNDRRLN--PEVSDTEAGSVSSGKASPTPN--FNGRKVSEYTVGVARE 342

```

; ORGANISM: Arabidopsis thaliana
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (596)..(598)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-60-752-355-51629

Query Match          7.0%; Score 186; DB 8; Length 884;
Best Local Similarity 24.7%; Pred. No. 4e-09;
Matches 84; Conservative 60; Mismatches 116; Indels 80; Gaps 17;

Oy 32 SPCRMGCRREQVNOGNMPELOS---TFLMADTEGEVYVWNE--LHFGRCKAFPAHEEK 86
Db 395 SPGRMRWK-----GRLLGMSGFGRHYLLGFNBSGEMCAMKVTLLCSDDPKS----RES 443
Oy 87 IQTVFEQVLVD--HPNIYKLHKRWLDITSEACARVIFT-TEYVSSGSLKQFLKTKGN 142
Db 444 AQQIGQRISTVLSRLRHQNIQVY----GSEYVDKXIYLEYVSGSIIYKLLQE----Y 494
Oy 143 KAMARAPAKMKWCQIIISALSFCAKSPPIIHGULTDPTIIONGLIKISVWHRIFSNA 202
Db 495 GQGEENAIRNYTQIIISGLAYLHAKN--TVHRIKGANIILVDHGRKVAIDFGMAKHITA 552
Oy 203 LRPEITALPDLLRSPIRAERELNHLHFPPEYGEVADGT--AVDIFSGCALEMAVLIEI 260
Db 553 QSGFLSKG---SP-----YMAPEVILKNSGSLTAVDINSGLCTVLEMAITTK 598
Oy 261 ----QTNGDTRVTEEARARHSLSDPNM-----REFITCLARDPAPRPSAHSLLF 308

```

QY 309 HRYLFVHSLKLLAHCFIOHOYIMPENVE-EKTKMDL 347

```

RESULT 15
US-60-752-355-18183
; Sequence 18183, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752.355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18183
; LENGTH: 883
; TYPE: PRY
; ORGANISM: Arabidopsis thaliana
US-60-752-355-18183

Query Match      6.9%; Score 185.5; DB 8; Length 883;
Best Local Similarity 24.8%; Pred. NO. 4,4e-09;
Matches 84; Conservative 60; Mismatches 116; Indels 79; Gaps 17;

```

```

Dh 395 SPGRMRKK-----GRLMGSPGHVYLLGFNSPBGEMCAMKEVTLCSDDPKS----RES 443
Oy 87 IQYFEBDLVLD--HNIYKLMKHYKMLDTSEACARVFT-TEYVSSSLKQFLKTKYKON 142
Dh 444 AQQGGQISVLSRLRHNIYOYY-----GSEYDDKKIYILEYVSSGGSIYKLLQE----Y 494
Oy 143 KAMARAMKMCCTQIISLTSFLHACSPPIHGNLTSPDTEIHNGLIKIGSVMHRIPSNA 202
Dh 495 GQFEMNIRMYTQIISGLYIAHKN--TYHRIKGNILVDPHGRKXVADFGMAKHITA 552
Oy 203 LRPELTLDLRSFIAREELRNHFFPEYGEVADGT--AVDIFSFGCALFEMAVLE- 255
Dh 553 QSGLSLFGK--SP-----YMAAEVINKNSNGSIAMVADWSLGTVELEMAATKPR 598

```

Qy 260 --1QTNQDTRVTEATARARSLSDPM-----REFILCCLAPDAPRPSAHSLTFH 309
 Db 599 PMSQYEG----VPAMFKIGNSKELPDIPDHLSEBEGDVFVKCLQRNPANRPTAQL-- 651
 Qy 310 RVLFEVHSLKLAHCFIOHOYLMPEVVE-EKTKAMD 347
 Db 652 -----DHAfVRNVMPMERPIVSGEPAPAMNV 677

Search completed: January 12, 2006, 11:30:50
 Job time : 50 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: January 12, 2006, 11:17:30 ; Search time 18 Seconds

(without alignments)
2710.105 million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670
Sequence: 1 MAAPBPAPRRARRERED.....RMKLAFLSTFLKYRGCTQA 507

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1346	50.4	258	2 T46491	hypothetical prote
2	543	20.3	231	2 T23136	hypothetical prote
3	415.5	15.6	461	2 T23574	hypothetical prote
4	372	13.9	516	2 T06716	hypothetical prote
5	368.5	13.8	1851	2 T19964	hypothetical prote
6	349.5	13.1	677	2 T02951	hypothetical prote
7	327.5	12.3	547	2 T46059	probable mitogen a
8	327.5	12.3	549	2 T51099	MAP kinase (improt
9	316.5	11.9	633	1 S49611	probable serine/th
10	195	7.3	1401	2 T39225	MAP kinase kinase
11	191	7.2	544	2 S40482	serine/threonine-s
12	190	7.1	366	2 T51339	mitogen-activated
13	190	7.1	345	2 G01773	p21-activated prot
14	188	7.0	525	2 S58682	protein kinase, p2
15	187.5	7.0	544	2 A57597	beta-p21-activated
16	186	7.0	378	2 T26684	hypothetical prote
17	185.5	6.9	544	2 T49376	p21 activated kina
18	185.5	6.9	883	2 A96662	hypothetical prote
19	185	6.9	738	2 F96701	hypothetical prote
20	183	6.9	1895	2 T06609	disease resistance
21	181.5	6.8	553	2 T01479	hypothetical prote
22	181	6.8	480	2 A86371	hypothetical prote
23	180.5	6.8	963	2 T09911	probable serine/th
24	177.5	6.6	342	2 T04054	protein kinase hom
25	177.5	6.6	736	2 T05132	protein kinase hom
26	176.5	6.6	403	2 JCS974	autora-related kin
27	175	6.6	481	2 T49072	protein kinase - m
28	174.5	6.5	608	2 G96575	probable MEK kinas
29	171	6.4	1008	2 D84434	probable receptor

30	170	6.4	372	2 T52621	mitogen-activated
31	170	6.4	391	2 T48115	protein kinase ATM
32	170	6.4	528	2 D86456	probable protein k
33	167.5	6.3	348	2 T51340	mitogen-activated
34	166.5	6.2	348	2 T52635	mitogen-activated
35	166.5	6.2	560	2 D85084	SNF1-related prote
36	166.5	6.2	746	2 S62365	NeK6 protein kinas
37	166	6.2	338	2 JCT838	probable Ste20-lik
38	164.5	6.2	487	2 A71407	hypothetical prote
39	163	6.1	1895	2 T15881	hypothetical prote
40	162.5	6.1	462	2 S29851	protein kinase 6 (
41	162.5	6.1	535	2 T51736	mitogen-activated
42	161.5	6.0	406	2 T52626	probable mitogen-a
43	161.5	6.0	706	2 A48084	SRF1 protein kina
44	161	6.0	722	2 T37970	probable G2-specif
45	160.5	6.0	677	2 T27127	hypothetical prote

ALIGNMENTS

RESULT 1

T46491 hypothetical protein DKFZp434P086.1 - human

C:Species: Homo sapiens (man)

C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

C:Accession: T46491

R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: 223037

A:Accession: T46491

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-258 <AAA>

A:Cross-references: UNIPROT:Q9NSY0; UNIPARC:UP1000006E116; EMBL:AL137662

A:Experimental source: adult testis; clone DKFZp434P086

C:Genetics:

A>Note: DKFZp434P086.1

Query Match 50.4%; Score 1346; DB 2; Length 258;
Best Local Similarity 100.0%; Pred. No. 2.6e-81;
Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	250	MCALENAVLEIQNGDTRVTEBAIARARHSLSDPNNEFTLLCCLARDPAPRPSAHSILFH	309
DB	1	MCALENAVLEIQNGDTRVTEBAIARARHSISDPNNEFTLLCCLARDPAPRPSAHSILFH	60
QY	310	RVLPEVHSLKLLAHGCFIOHOYLPENVVEKTKAMDHLAVLAELPPRRPPIQWRYSEV	369
DB	61	RVLPEVHSLKLLAHGCFIOHOYLPENVVEKTKAMDHLAVLAELPPRRPPIQWRYSEV	120
QY	370	SPFELDKFLDVNGVYPLNPNFATRPGLPRVLAPPPEEVQAKTPTPPSPDSETRKVI	429
DB	121	SPFELDKFLDVNGVYPLNPNFATRPGLPRVLAPPPEEVQAKTPTPPSPDSETRKVI	180
QY	430	OMOCNERSSDKARWHTLLVLVBDRILHROLTYDLLPTDQAOLASELVHYGFLHEDDM	489
DB	181	OMOCNERSSDKARWHTLLVLVBDRILHROLTYDLLPTDQAOLASELVHYGFLHEDDM	240
QY	490	KLAAPLESTFLKYRGCTQA 507	
DB	241	KLAAPLESTFLKYRGCTQA 258	

RESULT 2

T23136 hypothetical protein H37N21.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T23136; T23575

R:McMurray, A.

submitted to the EMBL Data Library, November 1997

A:Reference number: Z19694

A:Accession: T23136
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-231 <MTL>
A:Cross-references: UNIPROT:O45666; UNIPARC:UP10000076BE3; EMBL:AL008986; PIDD:CAA15621.
A:Experimental source: clone H37N21
R:McMurray, A.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19762
A:Accession: T23575
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-231 <MT2>
A:Cross-references: UNIPARC:UP10000076BE3; EMBL:Z75545; PIDD:CAA99888.1; GSPDB:GN00019;
A:Experimental source: clone K10D3
C:Genetics:
A:Gene: CSP:H37N21.1
A:Map position: 1
A:Introns: 31/3; 68/1; 122/3; 152/3; 171/3

Query Match 20.3%; Score 543; DB 2; Length 231;
Best Local Similarity 56.3%; Pred. No. 1,1e-28;
Matches 98; Conservative 31; Mismatches 45; Indels 0; Gaps 0;

QY 19 EDESEDESDILIESPCGRMQRREOVQGMPLQSTFLAMDTEBGEVYVNNELHFGDRK 78
DB 23 DSDADGAEELIESPDKRMKREQYKQRPIDAVAYLMDTENEVYVNNELHFGDRK 82
QY 79 AFAAHEEKIQTFEQLVVDHPNIVKHKYLDLTSACARVIFTEYVSSGSLKQFLKKT 138
DB 83 NFPAQBEKINAVPDNLTQLVHTLVKFKYWTDSKSEKPRITITEYMSGSMASFLQRT 142
QY 139 KKHKKAMNARAKMRCQTQILSALSFHACSPRIHNTLPTFIQNGIKIG 192
DB 143 RKAGSSLIAMKMKWTQILSALNYSDDPILHGLTGNVFIQNGIKIG 196

RESULT 3
T23574
hypothetical protein K10D3.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23574
R:McMurray, A.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19762
A:Accession: T23574
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-461 <MTL>
A:Cross-references: UNIPROT:Q21419; UNIPARC:UP10000076FE6; EMBL:Z75545; PIDD:CAA99887.1;
A:Experimental source: clone K10D3
C:Genetics:
A:Gene: CSP:K10D3.5
A:Map position: 1
A:Introns: 12/1; 40/1; 124/3; 157/3; 212/3; 345/2; 386/3; 421/2

Query Match 15.6%; Score 415.5; DB 2; Length 461;
Best Local Similarity 29.9%; Pred. No. 6.5e-20;
Matches 116; Conservative 57; Mismatches 92; Indels 123; Gaps 13;
QY 210 PDLLRPPIAREBELNLHFPPEYGEVADGT-AVLIFFSGMCALEMAV---LEITQND 265
DB 14 PDILNHVKTGRNMRYMAYIAPEYENNTELTSADYISFGISLEIAVIGLSGQONS 73
QY 266 TR--VTEBAIARAHSLSDPNMREFLICLADPARPSASHLFRVLFEVHSLKLLAA 323
DB 74 SEGPVTEVDYIEKAIRESLEDMDQDFTROCKRDPARPSARELLFFQILFEVHSLKLLAA 133
QY 324 HCFIIOQVLPENVVEKTKAMDHLAVLAEPLPRPRPPIQWRYSVSF--MELDKFLADV 381
DB 134 HAIVDSKKY--EDVSESAPRIKD-NETIATSKLR---EMAYCQVAAPQVDLEKFLDDV 186

QY 382 RNGIYPLANFA-----ATRP-----IG 398
DB 187 RNGIYPLTFAFLAHQPSITTLARXSTNPSITLITDIDASPSSTHPSANSTITAEISVNS 246
QY 399 LRPVLAP-----PEEVQK----- 412
DB 247 LFGQSQPSGCTTNTNGPSSIGKSASPEAVDKKIGEVSTESTSKVEVENGANVTIGSS 306
QY 413 ----AKTPPE-----PDSFTRKIQVQCNRESDEKARKMLTLLVLEDRILHQ 459
DB 307 NGRDAGSPTEEBEGEPNGRDLRLNRLILEINVLINENB-----MSIVLLEDMHQ 360
QY 460 LTYDLLPTSAQDLASELVHGFLEHD 487
DB 361 LTTISIKGNPPTLTENLITTHGFMQLD 388

RESULT 4
T06716
hypothetical protein T29H11.220 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-Oct-2004
C:Accession: T06716
R:Querter, F.; Cholene, N.; Robert, C.; Broctier, P.; Wincker, P.; Cactolico, L.; Artig
submitted to the Protein Sequence Database, April 1999
A:Reference number: Z15793
A:Accession: T06716
A:Molecule type: DNA
A:Residues: 1-516 <QNE>
A:Cross-references: UNIPROT:O9TK6; UNIPARC:UP100000A845D; EMBL:AL049659; GSPDB:GN00061
A:Experimental source: cultivar Columbia; BAC clone T29H11
C:Genetics:
A:Gene: ATSP:T29H11.220
A:Map position: 3
A:Introns: 23/3; 36/2; 112/2; 187/1; 239/3; 295/1; 338/1

Query Match 13.9%; Score 372; DB 2; Length 516;
Best Local Similarity 25.3%; Pred. No. 5.4e-17;
Matches 119; Conservative 88; Mismatches 176; Indels 88; Gaps 16;

QY 18 REDESEDESDILIESPCGRMQRREOVQGMPLQSTFLAMDTEBGEVYVNNELHFGDR 77
DB 2 RQDENNSEBEFVIDPTGRYGRYKVLGK--AFREYVARPDQLEGIVAMNQVRLDXK 58
QY 78 KAPAAHEEKIQTFEQLVVDHPNIVKHKYLDLTSACARVIFTEYVSSGSLKQFLKX 137
DB 59 FCSSEDLRLYSEVHLKTKLKHKSIIKFTSWIDHQ--MTINLTTEVPTSGNLKQYRK 116
QY 136 TKKNKAMNARAKMRCQTQILSALSFHACSPRIHNTLPTFIQNH-GLIKIG---- 192
DB 117 ----HKCVDLRLKMKSRDILGLVYLHSHDPVLRDLKCNIFNNGQGEVYKIDGL 172
QY 193 -SVWHRIFSNALRPPTALPDLLRSPPIAREBELRLHFPPEYGEVADGAVDIFSGMC 251
DB 173 AAILHR-----ARSHSVIGTPEFAPALYEDVDVAVDIYAFGC 213
QY 252 ALEMAVLEIQNGDT-----RTVEBAIARAHSLSDPNMREFLICLADPARPSAH 304
DB 214 LLELVTFEYRSECTNAQIYRKVTSIGIKPALLVWTDQVAFIEKCIAX-VSQRLSAK 272
QY 305 SLTFRVLFEVHSLKLLAHCFIOQVLPENVVEKTKAMDHLAVLAEPLPRPRPPIQW 364
DB 273 ELDDPFL-----KCYKEN-----TEVSSHKENGYNGLVDLSD----- 309
QY 365 RYSEVSPMELDKFLVDENGIYPLNFAATPLGLPRVLAPPEEYQAKTPTPEPDS 424
DB 310 --SEVGLTLVEQRDL-NTIFLKLRIDSK-----GQIRNHFPENIETDS 354
QY 425 TRKVIQMCNERSSEDKARWHLTLLVLEDRILQYTDLLFT---DSA 470
DB 355 FSVAIEMVEELDTDDO--ISTIAKMDITEISHIP-DWTSRLIGDOSA 402

QY 17 EREDESEDSIILLESSPCGRWQKRREOVNOCNMFGLOSTFLAMTDEGCVYVNNLHFGD 76

Db 2 EISSASDSISIAVETDPSPGRGFRFEVLGKGM---KTVYVAFQOVLGMEVANNQVLTNE 58

QY 77 RKAFAAHEKIQVTEOLVL---VDHPNIVKHKVLTDPSTACARVIFITEVSSGSLKO 133

Db 59 --VFPS-PEPIORLYSEVHLKKNLNHESIIRYCTSWIDVNRRTN--FTELEFTSGTLRE 113

QY 134 FLKTKTKKHKAMNAPAMRWCTOILSALSFHACSPPIIHGNTLSDTIFIO-HNGLIKIG 192

Db 114 YRRKTKQK---VDIAIKSWARQILNGLAYLHGHDPPVIRHDLKCDNI FVNGHLGQYKIG 169

QY 193 SVWHIFISNALRPPALPDDLRSPTRAREELRNLIH-----FPPPYGEVADGTAVDI 245

Db 170 DL-----GLAALIRGS-----QNAHSVIGTPEPMAPELVEEDYNEMLVDI 208

QY 246 FSPFGCAEMAVLEIQTMGDT-----RVTEEAIRARHSLSPNNREFILCCLADPA 298

Db 209 YSFGGCVLEMTLGEYPSYECNTPAQIYKKVTSKGLPDSFHLIQHTAQRFAFGKLT-BTVS 267

QY 299 RRPASHSLLFRVLFEVHSLKTLAHCFTIGQVLMPEVNEBEKRDANDLAVALAEFRPR 356

Db 268 RRLPAKELLADPFLAATDERDL--APLFRLLPQOLAIQTLAANGT-----VVEHLPSST 318

QY 359 RPPLOWRYSEVS-----FMEIDKFLKED--VANGIYPLMNEFAATRPGLPRLVA 404

Db 319 DPT---RTDMSINGKMSSEDTTIFLOQVILDGSGHMRNIQFP-----NILS 363

QY 405 PPPEVQAKATPTPEPPFDSSTRKVIQWQCNLERSEDKARWH-LTLLVLEDRHLRQULTYD 463

Db 364 DTPLELV-----ALEWKELEITD---MDPLEIAMIEHEISL----- 397

QY 464 LLPTDSADLASLSELVHYGFLHEDD 487

Db 398 LVPNMRAND--SSIRHESFGHEDD 419

RESULT 8

MS1099

MAP kinase [imported] - Arabidopsis thaliana

C|Species: Arabidopsis thaliana (mouse-ear cress)

C|Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004

C|Accession: T51099

R|Cooke, R.M.; Laudie, M.; Berger, C.; Delserny, M.

submitted to the EMBL Data Library, May 1999

A|Description: Analysis of Arabidopsis thaliana gene structure by cognate cDNA sequencing

A|Reference number: Z25297

A|Accession: T51099

A|Status: preliminary; translated from GB/EMBL/DBDJ

A|Molecule type: mRNA

A|Residues: 1-549 <COO>

A|Cross-references: UNIPROT:O9XFS5, UNIPARC:UP100000A62F3, EMBL:AJ238802, PIDD:CB43520

Query Match	12.3%	Score 327.5	DB 2	Length 549
Best Local Similarity	25.8%	Pred. No. 4.9e-14		
Matches 130	Conservative 79	Mismatches 176	Indels 119	Gaps 23
Qy	17 EREDESEBESDILIESPCGRWQKRREQVQCNMPGLQSTFLAMTEGEGEVVNNELHFGD	76		
Db	4 EISSASDSDSIAYVEHDPESGRGRPREVLGKAM---KTVYKAPFQVLYGEMVANNQVYKNE	60		
Qy	77 RKAPAAHEEKIQTYFEQLVL---VDHENIVYKLHKYMLDTSACARVIFITEVYSSGLKQ	133		
Db	61 --VPRF--PEPQRLRSEVHLKKNLHNSIIIRYCTSMIDVNNRTFN--FITELEPFGSTLRE	115		
Qy	134 FLKTKKXHKMANNAAMRWCTQIILSALSFACSPPIHGNLSDIFIQ--HGLKIKIG	192		
Db	116 YRRKQK---VDIRATSKMARQLINGLAVYHGHDPPVIRHDLKCDNIPLVNGHIGQYKIG	171		
Qy	193 SVMHRIFSNALRPPTALPDCLRSPRIABEREELRNLIH-----FPEPYEGVADGTAVDI	245		
Db	172 DL-----GLAALIRGS-----QNAHSVIGPPEMAELVEEDVNEVLVDI	210		

```

OY 246 FSEWCMLENAVLEIQNGDT-----RTBEVIAIARASISLDPNNKREITLCLIAADPA 298
Db 211 YSFEMCYLEMUTGSEYPSSECTNPAQIYKKATSGKLPDSFHLIOHTEAQRFVGKCL-ETVS 269
OY 299 RRPASHSLEFHRVLFEEVHSJLKLAAHCFOIHOYLMPENVVEBKTKAMDIAVLAELPRP 358
Db 270 RRLPAKLELADPFLAATVERDL-APLFRUPQALQNLANGT-----VVEHLPSTT 320
OY 359 RPLPQMKRYSEVS-----FMELDKFLEB-VRNGIYPLMFAATRPLGAPVLA 404
Db 321 DPL---TTTMSITGKNSSDHTIFLQVQILDDGGHNRNIOQPP-----NILS 365
OY 405 PPEEVOKAKTPPEPFDSERKVIOWOCNLSERSEDKARWH-LTLLVLIEDRLHRLQUTYD 463
Db 366 DTPELV-----ALEMVKLELITD---WDLPIAMIMENEISL----- 399
OY 464 LLPDSAQDLASELVHYGFLHEDD 487
Db 400 LVPMNRAND--SSIRHESFGCHEDD 421

```

```

RESULT 9
S49611
probable serine/threonine-specific protein kinase pkpa (EC 2.7.1.-) - Phycomyces blakesleyi
C|Species: Phycomyces blakesleyanus
C|Date: 10-Sep-1999 #sequence_revansion 10-Sep-1999 #text_change 10-Sep-1999
C|Accession: S49611; S59578
R|Rutiz-Perez, V., Murillo, F., Torres-Martinez, S.
submitted to the EMBL Data Library, November 1994
A|Description: A novel serine/threonine protein kinase in the fungus Phycomyces blakesleyi
A|Reference number: S49611
A|Accession: S49611
A|Molecule type: DNA
A|Residues: 1-633 <RUI>
A|Cross-references: UNIPARC:UPI0000172621; EMBL:Z46636
R|Rutiz-Perez, V.L.; Murillo, F.J.; Torres-Martinez, S.
Curr. Genet. 28, 309-316, 1995
A|Title: PKPA, a novel Phycomyces blakesleyanus serine/threonine protein kinase.
A|Reference number: S59578; MUID:96120859; PMID:8590476
A|Accession: S59578
A|Molecule type: DNA
A|Residues: 1-9/33-346 <RUW>
A|Cross-references: UNIPARC:UPI0000172622; UNIPARC:UPI0000172623; EMBL:Z46636
A|Note: only part of the coding region is given
C|Genetics:
A|Gene: pkpa
A|Introns: 60/1; 248/3; 277/2; 339/3; 534/1
C|Superfamily: Phycomyces blakesleyanus probable serine/threonine-specific protein kinase
C|Keywords: phosphotransferase; protein kinase
F|3/4-289/Domain: protein kinase homology <KIN>

Query Match      11.9%; Score 316.5; DB 1; Length 633;
Best Local Similarity 23.7%; Pred. No. 3.le-13;
Matches 126; Conservative 87; Mismatches 193; Indels 125; Gaps 23;

OY   EREREDSE--DESDILESPGCRWRQREOVACQNPGLOSTFLADTEEGEVVNNEL 72
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB   12 DMEVDVDSNMPDYEKVIEASGNRGYSKLTNVLDKG---AKVVYKAIDREAL----NDN 64
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY   73 HFGRKFAAAHEEKIQTVEQLVLVDHPNIIVKLHKTVLDTSEACARVIPTEVYSSGLK 132
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB   65 EITVTKVTROEFKDLGHEIDLKSVRHPIIITPHDAVYNETE---FVFITELMTSGTLR 120
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY   133 QELKTKKNKANNAARKWCOTIIISLSFLIACSPRIIHGULTPTTIO-HNGLIKI 191
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB   121 EYIRKLT--PLPIKIKVKRCRIQLIGLAIVLGHEPPIIHRDIKCNIIFINGAHGEIKI 177
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY   192 GSVMHRIFSNALRPPTALPDILSRPIAREEELRNLIFFPEE-YGEVAODGTAVDIFSFGM 250
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB   178 GDMGTAMRKGGKYTV-----IGTP-----EFMAPMYEEBGVNYEKVDIYAFGM 221
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY   251 CAEMAVLEIQTINGDT-----RVTEEARARASHSIDENKEFIICCLARDPPRRPSA 303

```

Db 222 CLEMAATGEYGESECTNANVQFKKVTOTIKRECSLRVODELLTVNICTLPED-ERM7A 280
 QY 304 HSLFHRVLE---EV-----HSLKLLAHNCF-----IQHY-----LM 333
 Db 281 QETLEHRRFLAVEBEVVLVSQDMTKLTLTLOVFKGDKLSVKTEFNADTDPAADVAMET 340
 QY 334 PENVVE---BKTAMDIAVLAELPRRPP-----LQMRSEVSFMEIDKFLIEDVRN 383
 Db 341 EBGVLQNCYQQLTCEINRLRDARNOQPPDKGDEDKIYWRENDRSLEERAKKDLAL 400
 QY 384 GIYPLAN-----FAATR-PLGLPRVLAPEEVOAKATPTPE 419
 Db 401 AVERVEAEKCKCELLBOHIIABERCKETIFALBOAKFQIPDLLOPQ---QPOPOPOQ 457
 QY 420 PPSERTRKVIQMCNLE-----RSEDKARHLLLVLEERLHR 458
 Db 458 P---QPOPOLOPOLQVLSPOSTSPGTPSTSDNST-NSTLSLSLELSK 504

RESULT 10

T39225
 MAP kinase kinase kinase - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 05-Oct-2004
 C/Accession: T39225
 R/Churher, C.M.; Gentles, S.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 Submitted to the EMBL Data Library, August 1997
 A/Reference number: Z21837
 A/Accession: T39225
 A/Status: preliminary; translated from GB/EMBL/DBD
 A/Molecule type: DNA
 A/Residues: 1-1401 <CHU>
 A/Cross-references: UNIPROT:O14299; UNIPARC:UPI0000138F13; EMBL:Z98763; PIDD:CA11500.1;
 A/Experimental source: strain 972h-; cosmid c961
 A/Genetic: SPDB:SPAC9G.02
 A/Map position: 1

Query Match 7.3%; Score 195; DB 2; Length 1401;
 Best Local Similarity 25.1%; Pred. No. 8.1e-05;
 Matches 87; Conservative 50; Mismatches 135; Indels 74; Gaps 18;

QY 36 RMQKRRQVQGMN--PGL-QSTFLAMDTEGSEVVNNEHFGDRKAFANHEKIQVFE 92
 Db 1036 RWO-----QGHFVSGMFGDYVTGVNMTGDLAVKEIKLQDSRTFRSTVQJHNEMT 1088
 QY 93 QLVLVDPHPI-----VLHKVWLDTSSEACARVIFITEVSSGSLKQFLKTKKXHKMANA 147
 Db 1089 VLRLNHPNVVTVYGVEVHR-----EKVYTFMEFCGGSLADL---AHGRIDE 1135
 QY 148 RANKRWCTQILSALFLHACSPPIHGNLTSDTIFIQHNGLIKIGSVMHRIFSNAT--RP 205
 Db 1136 NVLKVVYVQLLEGLAVYH--SQHILHRDIPANILDRGMKIKSD-----FSGALYVSP 1188
 QY 206 PTLAPDDLSPPIAREBELNHPPEY--GEVADT-----AVDIFSGMCALEMAV 257
 Db 1189 PTL-PEVRVEDIOPELOHLAG---TPMYAPPEIITLTKKGDGFAMDIWSLGVILEM-- 1241
 QY 258 LELQNGDTRVTEAL---ARARHSLSDPN-----MREPIICCLARDPARPSAHSLS 307
 Db 1242 MTGSTPSEMDNEKALHYVAAMHTPSIPONEKISSLARDPIEGCFERDEQRPRAVDL 1301
 QY 308 FHRVLFVHSLKLLAHCFIOHOYLMBNVNEEKTAMDLHVAVLAE 353
 Db 1302 THPWITDFRKKTIT-----MPATITKTSLS--HTIIEE 1335

RESULT 11

S40482
 serine/threonine-specific protein kinase (EC 2.7.1.-) - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 19-May-1994 #sequence_revision 10-Nov-1995 #text_change 05-Oct-2004
 C/Accession: S40482

R/Manser, E.; Leung, T.; Salihuddin, H.; Zhao, Z.; Lim, L.
 Nature 367, 40-46, 1994
 A/Title: A brain serine/threonine protein kinase activated by Cdc42 and Rac1.
 A/Reference number: S40482; MUID:94150588; PMID:8107774
 A/Accession: S40482
 A/Molecule type: mRNA
 A/Residues: 1-544 <MAN>
 A/Cross-references: UNIPARC:UPI000017A3E7
 C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 F/267-520/Domain: protein kinase homology <KIN>
 F/275-283/Region: protein kinase ATP-binding motif

Query Match 7.2%; Score 191; DB 2; Length 544;
 Best Local Similarity 24.6%; Pred. No. 4.4e-05;
 Matches 84; Conservative 60; Mismatches 115; Indels 82; Gaps 16;

QY 4 PEPAPRRAREREREDESESDILEE-----SPCGRWKRRBQVNGMPLQST 55
 Db 230 PPDALTRNTEKQKKPKMSDEE--ILEKLNIVSVGDPKKY--TRFEKIQG---ASGTV 283
 QY 56 FLAMDTEGSEVVNNEHFGDRKAFANHEKIQVFEQVLV--DHPNIVKLHKVWLD 112
 Db 284 YTMADVATGQEVAKIQMNL-----QQQPKKELINELVMEKNPNIVN---YLD 332
 QY 113 SEACARVIFITEVSSGSLKQFLKTKKXHKMANAAMKRWCTQILSALFLHACSPPI 172
 Db 333 YLVGDELWVMEYLAGSLTDVYET---CMBEQIAVACECLQALEFLH--SNQVI 385
 QY 173 HGNLTSDTIFIQHNGLIKIGSVMHRIFSNATRPPTALPDLRSPRAERELNHPF-- 230
 Db 386 HRDIKSDNILLGMDGVSKLDFG---FCAQI-----TPQSKSTVGTVPWMA 431
 QY 231 PEYGEVADTAVDIFSGCALEM-----AVLEIQNGDTRVTEALARA 276
 Db 432 PVVTRKAYGPKVDISLGMALTEGEBPYNLENPLRLYLIATNG---TPE----- 482
 QY 277 RHSLSDPN-----MREPIICCLARDPARPSAHSLSLHRLV 312
 Db 483 ---LQNPBKLSAIFRDLNRCLEMDVYKGSALKELOHPL 520

RESULT 12

T51339
 mitogen-activated protein kinase kinase (EC 2.7.1.-) 4 [validated] - Arabidopsis thaliana
 N/Alternate names: MAP kinase kinase 4
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
 A/Accession: T51339
 A/Status: preliminary; translated from GB/EMBL/DBD
 A/Molecule type: mRNA
 A/Residues: 1-366 <ICH>
 A/Cross-references: UNIPROT:O80397; UNIPARC:UPI000009FE2F; EMBL:AB015315; PIDD:BAA28830
 A/Experimental source: strain Columbia
 C/Genetic: A/Genetic: ATMKK4
 C/Function: A/Description: (EC 2.7.1.-); mitogen-activated protein kinase kinase [validated, MUID:9
 C/Superfamily: kinase-related transforming protein; protein kinase homology
 C/Keywords: phosphotransferase; protein kinase

Query Match 7.1%; Score 190; DB 2; Length 366;
 Best Local Similarity 22.3%; Pred. No. 3.1e-05;
 Matches 81; Conservative 54; Mismatches 122; Indels 106; Gaps 16;

QY 45 NOGNMPLQSTFLAMDTEGSEVVNNEHFGDRKAFANHEKIQVFEQVLV--DHPNIVKLHKVWLD 112
 Db 51 SSGSAPSSSGSASTTNSIEAKVNSDLYRGNRIGSGAGCTYVYKVIHRSSRLYALKVI 110

```

OY      83  ---HEKIO---TYEOLVVDHPIVYKHLKMYLDJEBACRNVFIIEYVSSGLKQFL 135
         | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      111 YGNHEETVARQICREIEILRDVNHHPVVKCHEMFQNGE---IOVLBFMDKSGLE--- 163
         | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      136 KTKXKNHKAANAWKR-----WCTQILSLSFLHACSPRIIGHNLTSDTTFIOHNGLIK 190
         | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      164 -----GHVWKEQOLADLSROIISGLAYLH--SRHIVHRDIKPSNLLINAKRYK 211
         | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      191 IGSVA-HRIFSNALRPPTALPDDLRSPIRAERERLNIHFPRP-----YGEVADGTA 242
         | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      212 IADFEVSKILAQTMPCNS-----SVGTIAYMSPEKINDLDNGKT-DGYA 256
         | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      243 VDISEFGCALEMAV---LEIQINGDTRVTEBATAIAR---HSLSDPNNEFTLCTIA 294
         | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      257 GDIMWLGVSILFEFYIGRPFPPVSRQGDWASLMCAICMSQPREAPRTASPEFHNFIISCIO 316
         | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      295 RDPARPRSAHSLLFRVLFEVYSLKILANCFIOHQYLMREKVVAEKKRAMDLIVLDEL 354
         | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      317 REFGRRRSAMOLL-----QHPTLRASPSQNRK--PQNLHQLRPP- 354
         | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      355 PRP 357
         | | |
Db      355 PRP 357

```

```

RESULT 13
p21-activated protein kinase - human
G01773
C|Species: Homo sapiens (man)
C|Dates: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 05-Oct-2004
C|Accession: G01773
R|Chernoff, J
submitted to the EMBL Data Library, April 1995
A|Reference number: G08374
A|Accession: G01773
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: mRNA
A|Residues: 1-545 <CHE>
A|Cross-references: UNIPROT:013153, UNIPARC:UPI00000355B3, EMBL:U24152, NID:g780805, PIR:
F|Gene: Pak1
F|268-521/Domain: protein kinase homology <KIN>

Query Match          7.1%  Score 190;  DB 2;  Length 545;
Best Local Similarity 24.6%  Pred. No.5,1e-05;
Matches 84;  Conservative 60;  Mismatches 115;  Indels 82;  Gaps 16;

QY  4 PEPAPRRAREREREDESEDSILIE-----SPGGRWQKRRQVNOGNMFGLOST 55
DB  231 PPDALTRTEKQKKPKMSDEE--ILEKLRISIVSGDPKKKY-TRFEKIGG---ASGTV 284

QY  56 FLADTTEGCVVNMELHFGDKRKAFAAEKIQIVFEQVLV--DHFNIVKLAHKYIDT 112
DB  285 YTADNVATGQEAIAKKMVL-----QQQPKKELINILWRKNKNIVN---YIDS 333

QY  113 SEACARVIFITEVSSGSLKQFLKTKKKNHAKANARAKWCTOILSFLHACSPPI 172
DB  334 YLVGDELWVMEYLAGSSLTDVYET-----CMDSGQIAVCREGLQLEFLH--SNQVI 386

QY  173 HGNLTSDTIFIQHGLIKIGSVWHRIFSNALRPPTALPDDLRSPIRAERELRNLFHP-- 230
DB  387 HRDIKSDIILLGMSGSVKLTDPFG--FCAQI-----TPQSKRSFMVGTPYVMA 432

QY  231 PPEYGEVADGNAVHIFSRGMCALEM-----AVLEIQTNDDTVTVEAARAA 276
DB  433 PEVVTTRKAYGPKVDWISGIMAIEMIEEGEPPLYNENPIRALYTLATNG---TPE----- 483

QY  277 RHSLSDPN---MREFLCCIARDPARPARSAHSLLFRRVL 312
DB  484 ---LQNPPEKLSAIPRDPLNRCIDMDVVEKRGSAKELQHQPL 521

RESULT 14

```

[illegible]

QY 270 BEAIAARHSLSDPNMREFLICCLARDPARPSAHSILFHRVL 312
 Db 466 NP-----EKLTS-PIFPDPLNRCLEMDVEKRGSAKELLQHPFL 501

RESULT 15

A57597

beta-p21-activated protein kinase - rat

N/Alternate names: beta-PAK

C/Species: Rattus norvegicus (Norway rat)

C/Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 31-Dec-2004

C/Accession: A57597

R/Manser, E.; Chong, C.; Zhao, Z.S.; Leung, T.; Michael, G.; Hall, C.; Lim, L.

J. Biol. Chem. 270, 25070-25078, 1995

A/Title: Molecular cloning of a new member of the p21-Cdc42/Rac-activated kinase (PAK)

A/Reference number: A57597; MUID:96027610; PMID:7559638

A/Accession: A57597

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-544 <MAN>

A/Cross-references: UNIPROT:Q62828; UNIPARC:UPI0000131225; GB:U33314; NID:G1039424; PION

F;266-519/Domain: protein kinase homology <KIN>

F;274-282/Region: protein kinase ATP-binding motif

Query Match 7.0%; Score 187.5; DB 2; Length 544;

Best Local Similarity 24.6%; Pred. No. 7.4e-05;

Matches 87; Conservative 57; Mismatches 119; Indels 91; Gaps 17;

QY 2 AAD-EEAPARRA-----REREREDESEDESILRE-----SPCGRWQKRE 42
 Db 214 AAKNKETPPSAENANSTLYRNTDRQKSKMTDEILEKRSIVSGDPKKY-TRFE 272
 QY 43 QVNOGNMPSGLQSTFLANDTEEGVEVWNNEHFGDRKAFAAHEKIQTFEQVLV---DH 99
 Db 273 KIQGG---ASGTVYALDIDATGOEVAIKQNNL-----QQPKKELINEILVMRENK 322
 QY 100 PNTVKLHKYLDTSSEACARYIFITEVYSSGSLKQFLKKTGKHAKANARAKWCTQILS 159
 Db 323 PNTVN---YLDYLVGDELVMVMEYLAGSLTDVYTET-----CMDEGQIAAVCRECLQ 373
 QY 160 ALSFLHACSPPIIHGNLSDTIFIOHNGLIKIGSVWHRIPSNALRPPTALPDDLRSPIRA 219
 Db 374 ALDPLH--SNQVIHRDIDKSDNIIIGMDGSVKLTDFG--FCAOI-----TPEOS 417
 QY 220 EREELRNLPF--PPEYGEVADGTAVDIFSGMCALEM-----AVLBIOYN 263
 Db 418 KRSTWGTPTWMAPEVYTRKAYGPKVDIMGLIMATMEVGEPPYLNNENPLRALYLIATN 477
 QY 264 GDRFVTEEAIAARHSLSDPN-----MREFLICCLARDPARPSAHSILFHRVL 312
 Db 478 G---TPF-----LQNPRLSAVFRDPLNRCLEMDVDRGSAKELLQHPFL 519

Search completed: January 12, 2006, 11:22:27
 Job time : 20 secs

THIS PAGE BLANK (USPTO)